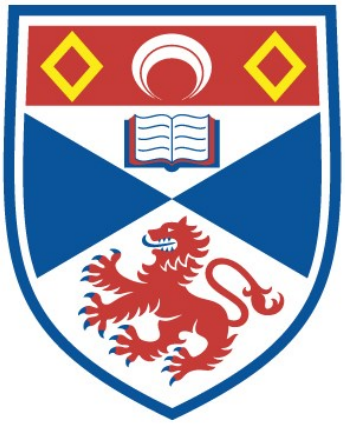
THE NEUROPHYSIOLOGY OF FORM AND MOTION PROCESSING IN THE TEMPORAL LOBE OF THE MACAQUE MONKEY

Mike W. Oram

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



1995

THE NEUROPHYSIOLOGY OF FORM AND MOTION PROCESSING IN THE TEMPORAL LOBE OF THE MACAQUE MONKEY

Mike W. Oram

School of Psychology,

University of St Andrews,

Fife, KY16 9JU, U.K.

DECLARATION FOR THE DEGREE OF Ph.D.

I, Michael William Oram, hereby certify that this thesis has been composed by myself, that it is a record of my own work, and that it has not been accepted in partial or complete fulfilment of any other degree or professional qualification.

Signed..... Date..27/4/95

I was admitted to the Faculty of Science of the University of St. Andrews under Ordinance General No. 12 on 1st October 1988 and as a candidate for the degree in October 1990.

Signed..... Date...27/4/95

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate to the degree of Ph.D.

Signature of Supervisor..... Date.....

THESIS COPYRIGHT DECLARATION

Unrestricted
In submitting this thesis to the University of St. Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any *bona fide* library or research worker.

Signed..... Date.27/4/95

ACKNOWLEDGEMENTS

The greatest debt is undoubtedly to Dave Perrett. His ability to give me almost enough rope to hang myself, but stop me short of putting the noose around my neck gave me the encouragement and confidence to believe in myself as a researcher. Dave's enjoyment of his work has given me the most important lesson of all: enjoying work means that its not really work. Verity Brown, Eric Bowman, Peter Foldiak and Dave Carey provided encouragement and useful discussion.

My fellow workers and conspirators (especially Jari Hietanen, Mark Harries, Phil Benson, Sharon Thomas, Nathan Emery and Lisi Wachsmuth) distracted me from the day to day repetitiveness and reminded me to laugh at the world. A special thanks to Ann Perrett who was always willing to proof read manuscripts and correct the numerous errors.

The "upstairs" technicians (Hilda Dickie, Wendy Taylor and Derek Brook) made the day to day running of the laboratory simple and easy. Its a task for which they never get the thanks that they deserve. A special thanks also to Mary Latimer for all the histological help and supplying high-quality electrodes.

The "downstairs" technicians (Malcom McClandless, Pete, Brian Kirk, Andy Burnley, Rab Gay, Jackie, Ken, to name but a few) were always understanding and grasped the idea of what was needed with minimal effort. There expertise is second to none, and without there ability much of the work presented here would not have been possible.

To the bar staff of my regular haunts (Malcom, Ellis, the two Richards, Connie, Ian, John, Gavin, Willie, Clare, and Alison) and drinking companions (Tim, Hannah, Dave, Paul, Duncan, Russ, Adam, Bill, Geoff, Beth, Ronnie, Stuart) thanks for helping me relax and unwind and then listening to my drunken ramblings. And finally to Frances Marsh, for all of the above and much much more.

CONTENTS

ABSTRACT

The neurophysiology of form and motion processing in the temporal lobe of the macaque monkey

M.W.Oram, School of Psychology, St. Andrews University.

Consideration of available evidence suggests that primate vision utilises two parallel cortical pathways to process visual information. The ventral pathway processes form or shape information, while the dorsal pathway processes motion information. In the macaque monkey, the superior temporal sulcus in the temporal lobe is one of the few cortical areas that receives input from both these pathways.

In this thesis recordings from visually responsive neurons in the macaque superior temporal sulcus are described. The cell response properties of three cell groups are investigated. One cell population show selectivity for the sight of static images of particular views of the body. The second group of cells shows selectivity for the sight of objects moving in the environment, independent of the object's form. The final group of cells show selectivity for particular views of the body providing that they are moving in particular directions.

The responses from these three groups of cell types are subjected to an analysis technique that allows insights into possible computational processes underlying the observed neural selectivities.

In particular, it is argued that the primate visual system processes form information primarily in a feedforward way, a property few computational models of visual processing employ. These data are combined with data from other studies to produce a speculative outline for a biologically plausible model of primate visual form processing.

The recordings also revealed cell responses to walking bodies that showed a remarkable selectivity for "structure from motion". It is suggested that this selectivity is developed by associative learning between the initially separate form and motion inputs.

Investigation of the integration of form and motion information onto single cells indicated a hitherto unforeseen problem: a temporal asynchrony between the arrival times of form and direction information. This asynchrony indicates that previously proposed mechanisms for binding of information about the same object are incorrect.

CHAPTER 1

INTRODUCTION

(Oram & Perrett, Neural Networks, 7, 945-972, 1994)

1. OVERVIEW

One of the most investigated topics of primate cortical function is that of vision. Neurobiological data concerning static visual image processing is available from anatomical, lesion, neurophysiological and event related potential (ERP) studies in monkeys and from ERP, positron emission topography (PET) and neuropsychological case studies in humans. The wealth of data available has highlighted a startling fact: we know very little about how visual processing leads to perception! A plethora of processing models have been proposed, yet few encompass or account adequately for more than a small fraction of the available neurobiological data. The computational power of the visual system is indicated by the failures of such models. Due to this computational complexity of visual processing, it seems reasonable to suppose that our understanding of the biological basis of vision will only come about from knowledge of the precise computational processes that underlie the selectivities of the cells within the visual system: in other words, understanding will come with modelling.

The work presented here does little (if anything) to change the situation of the inadequacy of current computational approaches, except pose constraints on the possible classes of models that can be used to account for the neurophysiological data. The constraints are derived from investigations into the processing efficiency of the macaque visual system. The data come from examination of the two major pathways in primate visual processing (the form

and the motion pathways) and also from examination of the integration of these two types of visual information.

2. THE VENTRAL PATHWAY

The purpose of this introduction is to review some of the constraints on biological vision of static shape (form) and to synthesize a speculative model that reflects current knowledge of neurobiology. A review of the motion processing pathway is given in chapters 5 and 6. To put this present review of the neurophysiology of static visual form processing into a more general framework, it is first necessary to describe the basic anatomical structure of the macaque cortical visual system (see Figures 1.1-1.2). Figure 1.3 summarizes the flow of information through what are considered to be the most important brain areas for understanding object recognition. Following these anatomical details, brief descriptions of the stimulus selectivities of cells (neurons) are given for each of the anatomical regions. Figure 1.6 summarizes the type of selectivity found in each cortical area. From the descriptions of cell response selectivities an overall picture or framework of the neurophysiology of static visual form processing is developed. Having described the basic properties of the primate visual system, brief consideration is given to the neurophysiological evidence suggesting that expectation has an important role in high level visual information processing (as well as other modalities). It is suggested that the columnar organisation of cerebral cortex allows coding of the particular attributes of a pattern by their distinctiveness relative to the norm of each attribute class. Together, these considerations are integrated to provide a summary of the constraints operating on visual information processing in the primate visual system and an outline of a biologically plausible model of visual processing underlying object recognition.

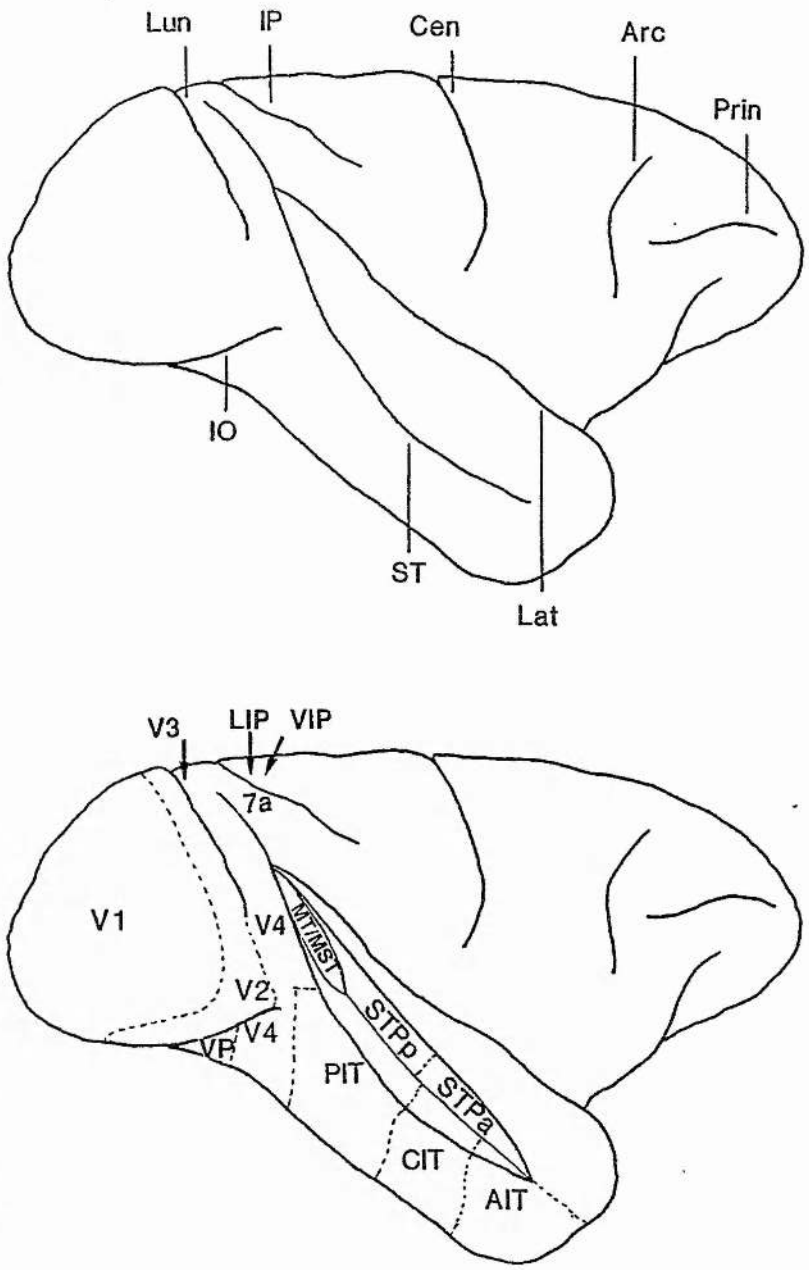


FIGURE 1.1. VISUAL CORTICAL AREAS OF THE MACAQUE BRAIN. Schematic view of the right side of the Rhesus macaque brain. UPPER: THE MAJOR SULCI. Moving from posterior (left) to anterior (right), the lunate (Lun) and inferior occipital (IO) sulci mark the anterior limits of the occipital lobe. Above the lateral sulcus (Lat) and posterior to the central sulcus (Cen) lies the parietal lobe with the intra-parietal sulcus (IP) running in the anterio-lateral direction. Anterior to the central sulcus is the frontal lobe with the arcuate (Arc) and principal (Prin) sulci. Below the lateral sulcus lies the temporal lobe, with the superior temporal sulcus (ST) running along its length. LOWER: THE VISUALLY RESPONSIVE CORTICAL AREAS OF THE OCCIPITAL, PARIETAL AND TEMPORAL LOBES. The occipital lobe is dominated by the primary visual area (V1). The second visual area (V2) forms a strip running along its anterior part. V3 lies in the bottom of the lunate sulcus bordered by V2 and V4. In the parietal lobes, area 7a lies on the lateral surface, with the lateral and ventral intra-parietal areas (LIP and VIP) within the intra-parietal sulcus. The superior temporal sulcus has been opened out to show the medial and medial superior temporal areas (MT and MST respectively) at the posterior end and, running anterior, the posterior and anterior superior temporal polysensory areas (STPp and STPa). Below the superior temporal sulcus, along the lateral surface of the temporal lobe lie the posterior, central and anterior infero-temporal areas (PIT, CIT and AIT respectively).

3. THE ANATOMY OF THE MACAQUE VISUAL SYSTEM

This section provides a overview of the visual anatomy. While the precise anatomical details are used as a basis of arguments about the necessary sequence and number of computations (chapters 3 and 5), the essential information is summarized in Figure 1.3. Therefore, readers more interested in the nature of the visual computations carried out should be able to understand subsequent sections on the basis of the summary Figures 1.1-1.3.

Figure 1.1 shows a schematic representation of the right side of the macaque cerebral cortex. In the upper figure the major sulci (folds in the cortex) are marked, while in the lower figure the major cortical areas containing visually responsive neurons are shown. The pattern of connections between these (and other) brain areas is complex. With further subdivisions there are more than 35 anatomically separate cortical areas which contain visually responsive neurons. The pattern of connections between these areas is not as daunting as it might first seem. Having correlated most of the available anatomical data available on the macaque visual system, Felleman and Van Essen (1991) noted that the connectivity between the areas was not as complete as it could have been (i.e. it was not the case that every area was connected to every other area). Broadly speaking, Felleman and Van Essen argued that the visual system could be divided into two main processing streams; a suggestion that has also been made from consideration of neuropsychological deficits and neurophysiological studies (Newcombe and Russell 1965; Ungerleider and Mishkin 1982; Mishkin et al. 1983; De Yoe and Van Essen 1988; Maunsell and Newsome 1987; Merigan and Maunsell 1993). This anatomical separation into two streams has also been shown using formal analysis of the number of anatomical connections between visually responsive cortical areas (Young 1992).

In primates the vast majority of visual information arrives in the cortex via the lateral geniculate nucleus (LGN, a subcortical structure) to area V1 in the occipital lobe at the back of the brain. Connections also exist from the LGN to prestriate areas V2, V3/V3a, V4 and MT (Benevento and Yoshida 1981; Bullier and Kennedy 1983; Fries 1981; Wong-Riley 1976) but not to the more anterior regions of the temporal lobe (STPp or STPa, Fries 1981; Iwai et al. 1980; Yukie and Iwai 1981). Visual information is also passed to these visually responsive cortical regions via the pulvinar from the superior colliculus (Bruce et al. 1986; Girard and Bullier 1989; Finlay et al. 1976; Girard et al. 1991; Gross 1991; Rodman et al. 1989, 1990; Mizuno et al. 1981). Neurophysiological studies of the areas following selective lesions and cortical cooling of V1 (thus inactivating area V1) indicate, however, that the direct geniculate and pulvinar connections to these prestriate areas are substantially weaker in terms of influence compared to those from the LGN via V1 (Bruce et al. 1986; Rodman et al. 1989, 1990). Despite these connections being weak, connections to visual cortical areas other than V1 may have a functional role (Gross 1991) which become apparent under extreme conditions such as those following brain damage.

Details of the complexity of the connections within visual cortical areas are outside the scope of this article. The main feature of interest here is that cortical areas are six layered structures and have distinct input and output layers: the input to cortical regions is predominantly to layer IV, whereas the output leaves predominantly from infra-granular layer VI and the supra-granular layers (I-III) (Felleman and Van Essen 1991). Thus for simplicity each cortical area can be considered as having a minimum of two functional layers: an input and an output layer. This is not to say that there are only two layers involved in processing: indeed all cortical areas have 6 distinct functionally active layers, and in many of these layers (particularly in V1) have further functional sub-divisions. The vast majority of connections between cortical areas are reciprocal: if one area projects to a second, then it is very likely that the second area has a projection

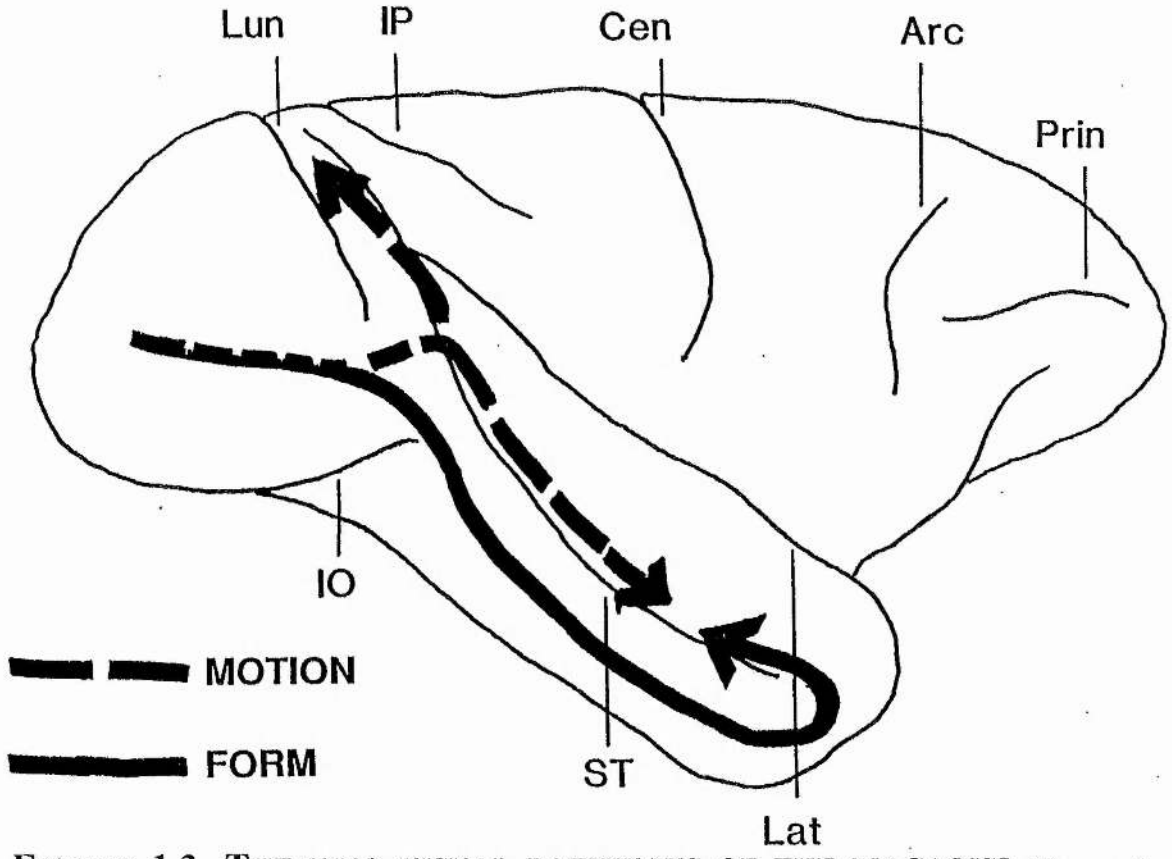


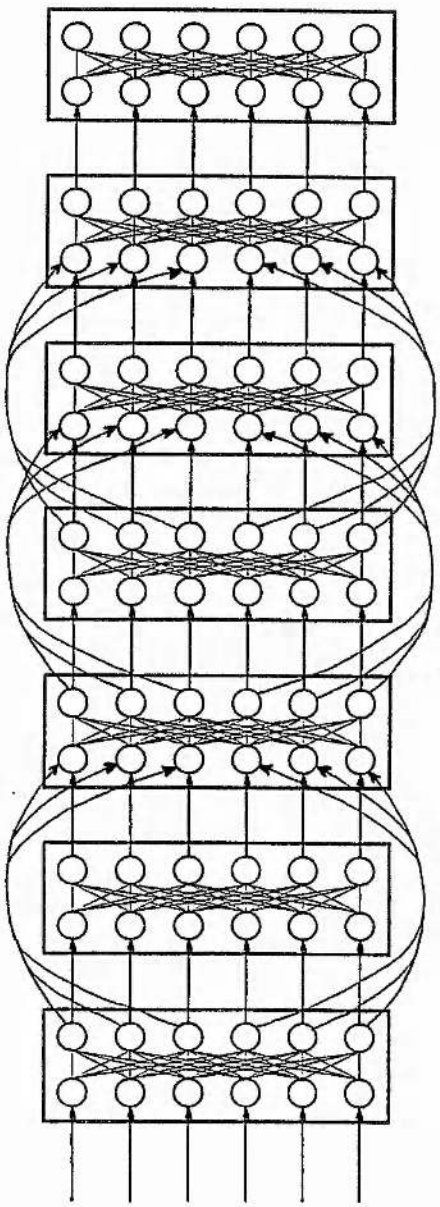
FIGURE 1.2. THE TWO VISUAL PATHWAYS OF THE MACAQUE. The motion pathway (also called the dorsal or 'where?' pathway) is shown in the thick dashed line superimposed on the right view of the macaque brain. The form pathway (also called the ventral or 'what?' pathway) is shown as the solid thick line. [Abbreviations of sulci as in Figure 1.1 upper].

back to the first (these feed-back connections can originate from either layers I-III or VI).

Zeki (1973, 1977) was first to show that cells in the anatomically distinct regions showed response selectivities for different stimulus attributes (for review see Zeki 1974b, 1990). Evidence from subsequent neuropsychological, neurophysiological and lesion data suggests a remarkable division of function in primate visual processing: one sub-system for motion processing and another for form or shape processing. [Zeki has argued that visual processing occurs along 3 pathways: form, colour and motion (Zeki and Shipp 1988; Zeki 1990).] This segregation of function follows the idea of the anatomical separation into two pathways. The pathways have variously been called the 'what?' and 'where?' pathway (Ungerleider and Mishkin 1982) or the ventral and dorsal pathway (De Yoe and Van Essen 1988). More recent consideration of neuropsychological case study data has led to the suggestion that the dorsal or 'where?' pathway could perhaps be more accurately labelled as the 'how?' pathway since it may involve control of motor actions in relation to visual information (Goodale and Milner 1992; Milner and Goodale 1993; Goodale 1993). Figure 1.2 shows the general division of these two pathways. The dorsal stream passes from V1 to V2 and MT. Cells from MT project to MST and FST as well as LIP and VIP. Within MST/FST there is a splitting of the dorsal pathway, with connections both to the parietal areas and to temporal areas through projections along the superior temporal sulcus to the posterior and anterior superior polysensory areas (STPp and STPa). A more detailed review of the dorsal or motion pathway is provided in chapters 5,6 and 7.

Although there is slight disagreement on the functional interpretation of the dorsal stream of processing, virtually all investigators classify the ventral stream as a visual pattern processing system. The ventral pathway passes from V1 to V2-V4 and from there along the inferotemporal cortex (through the posterior, central and anterior areas PIT, CIT and AIT) to STPa. Figure 1.3 shows

FIGURE 1.3. SCHEMATIC REPRESENTATION OF THE ARCHITECTURE OF THE FORM PATHWAY. The major visual areas comprising the form pathway are shown [abbreviations as in Figure 1.1 lower]. Each area is represented as a box with an input and output layer of cells. Within each area there is considerable cross-connection between elements. The maintenance of retinotopic mapping in areas V2/V3 and V4 implies that there is one to one mapping between these areas. This has been maintained in the figure for all connections between the cortical areas as there is no evidence for a greater spread in connections between subsequent areas. Reciprocal connections between cortical areas are not shown for simplicity.



schematically the major visual areas found along the ventral pathway with the possible jumps in the hierarchy shown. The reciprocal feed-back connections between cortical areas have not been shown. Each area has been assigned two layers corresponding to the input and output layers.

The areas along the ventral pathway provide some reciprocal connections to sub-cortical areas including the tail of caudate, the claustrum, the amygdala and the hippocampal complex (Yeterian and van Hosen 1978; Webster et al. 1993; Baizer et al. 1993; Steele and Weller 1993). Interestingly, parietal cortex in the dorsal pathway projects to similar areas but the particular termination sites within these sub-cortical areas suggest that the separation of the two pathways is maintained (Baizer et al. 1993). These sub-cortical areas may play a role in a variety of 'higher level' or more 'cognitive' functions, such as habit formation, emotional states, associative memory recall and recognition (Brothers and Ring 1992, 1993; Wilson and Rolls 1993; Cahusac et al. 1991; Ono et al. 1991; Squire et al. 1989, 1990; Squire 1992).

As an aside, it is noted here that these sub-cortical connections indicate that information may not have to pass to the end of the visual processing chain (AIT/STPa) before it can be passed to these sub-cortical areas. As a consequence, evidence for object recognition may be accumulated from elementary object features (e.g. it is yellow so it might be a lemon but it is not an orange). It is possible therefore that these anatomical pathways might underlie observations that a number of different visual attributes can be used to support object recognition (Horel 1993). This type of evidence accumulation could be equated to cascade processing proposed within 'black box' psychological models (Humphrey et al. 1988). Here it is suggested that each 'stage' of the cascade could be mediated by separate cortical regions. It seems likely that utilization of information available from earlier within the cascade depends on the task. If a subject is asked "Is this a picture of Marylin Monroe or John F. Kennedy?" the solution is available from information resolved at an early stage of processing "Is the top of the picture

blonde or dark?". Therefore the answer to this question is available without the need to know what sort of object or facial configuration is being viewed. Under normal circumstances, however, the question "What is this a picture of?" is rarely soluble in this way: the whole image has to be processed before an answer can be given. It is this latter processing that is considered in this article.

4. THE NEUROPHYSIOLOGY OF VISUAL PATTERN PROCESSING

Comparison of Figures 1.1 and 1.2 shows that the form processing stream passes through areas V1, V2, V3, V4, PIT, CIT and AIT and into area STPa in the superior temporal sulcus. Examination of the termination layers of connections between these areas indicates that there is a hierarchy of the areas with V1 at the lower end and areas such as STPa forming the later areas. Other 'high level' cortical areas (e.g. entorhinal cortex of the hippocampal complex, the temporal pole and the ventral surface of the temporal lobe) are absent from this scheme for simplicity. These other areas may be equally distant from the retinal input but their visually responsive cells have yet to be studied in the same detail (though see Miyashita and Chang 1988; Sakai and Miyashita 1991; Li et al. 1993; Miller et al. 1993; Fahy et al. 1993). The direction of the proposed hierarchy passes from posterior to anterior regions as indicated by the arrowheads in Figures 1.2 and 1.3. The stimulus selectivities shown by cells in progressively higher areas and their possible computational function are considered below.

Contour Extraction

The neural responses in V1 are typically selective for orientation and spatial scale (Hubel and Wiesel 1968; De Valois 1982a,b), for example a horizontal, thin bar not a horizontal, thick bar. [Other properties of V1 cells are considered below.] Cells in V1 do not respond to stimuli in all parts of visual

space, but rather to stimuli in specific regions called the classical receptive field. The classical receptive fields of V1 cells are very small in diameter (e.g. median of 0.1 degree^2 area at eccentricities of less than 5 degrees, Lehky et al. 1992) and rarely above 3 degree^2 area at greater eccentricities (although Schiller et al. 1976a report V1 complex cell classical receptive field sizes of up to 3 degree^2 within 5 degrees of the fovea). There are two basic classes of stimulus selectivity in V1: cells whose responses are phase dependent (i.e. selective for a bright line or bar on a dark background but not a dark line on a light background - simple cells) and cells whose responses are phase independent (a line defined by contrast in either direction, light on dark or dark on light) the complex cells. More strictly speaking the responses of complex cells is non-linear with respect to inputs within the receptive field, whereas for simple cells the responses are a linear combination of these inputs. There is a third group of cells, often called hypercomplex or end-stopped cells, which are not only selective for orientation but also for line length (e.g. Hubel and Wiesel 1968; De Valois 1982a,b; Schiller et al. 1976a). The sensitivity to line length can be present for one or both ends of the stimulus line.

Originally it was thought that the cells of V1 were only influenced by stimuli presented within the classical receptive field which can be mapped using flashing bars edges and dots. This no longer seems so clear, as more recent evidence suggests that some cells in V1 are also sensitive to influences of the surround: Knierim and Van Essen (1992, see also Allman et al. 1985) have shown that as many as 80% of V1 cell responses can be modulated by a textured surround falling outside the classic receptive field. For most cells the response was stronger when the elements in the surround were orthogonally aligned relative to the preferred stimulus orientation in the classical receptive field. This modulation of responses by the surround could in part underlie the perceptual 'pop-out' effects where easy to see contours are defined by texture difference (Knierim and Van Essen 1992, see also Fiorani et al. 1992).

The classic receptive fields of cells of V1 are not arranged in a random order: there is a clear retinotopic map present, such that adjacent cells have adjacent receptive field positions in the retina. Orientation selectivity also has a non-random arrangement: there are columns passing through the layers of the cortex containing cells all with similar orientation preferences. Each local area of the retinotopic map contains enough columns to cover all orientation selectivities (Hubel and Wiesel 1968; Blasdel et al 1992a,b). Furthermore, those cells of V1 which are selective for the direction of stimulus motion are clumped together, as are those which respond independent of motion. Cells selective for stimulus colour are also clumped and separate from cells selective only for luminance contrast or only for direction of motion (Hubel and Wiesel 1968; Michael 1981). Finally, the neural response sensitivity to which eye received the visual input (ocular dominance) is also highly organised: within layer IV (and adjacent layers) units are driven best by input from one eye or another such that there is an alternating pattern of left/right eye dominance columns. The more superficial and deeper cortical layers show a less clear pattern of ocular dominance and more cells are driven by input from either eye (Hubel and Wiesel 1968; Blasdel et al. 1992a,b; see Muly and Fitzpatrick 1992 for detailed anatomical study of the convergence of monocular to binocular sensitive cells). Further, a greater proportion of complex cells show this binocular sensitivity than simple cells (Schiller et al. 1976b). The concept of regular columnar organization is not, however, universally accepted, especially the notion of regular sized, repeating columns (e.g. Swindale 1990).

From the early work of Hubel and Wiesel (1968) it seemed that cells in V1 were detecting 'lines' and 'edges' of visual objects. As methods of stimulus presentation became more controlled and mathematically defined, researchers characterized the spatial frequency and contrast sensitivity of these cells (e.g. Campbell and Robson 1968; Schiller et al. 1976c; De Valois et al. 1982a,b). This shift in emphasis to investigations of spatial frequency selectivity made people

realize that detection of an edge of a real object within an image would require integration of information over many spatial scales. This approach, however, may have led to the erroneous assumption that analysis of spatial frequency coding *is* visual processing. The analysis of *local* variation in light intensity could well be an important function of early visual processing in order to define object boundaries. Gabor filter shape of V1 receptive fields can be regarded as representing a compromise, allowing for detection of spatial frequencies (periodicity) over extended image regions (the cosine component) and the position of a local feature (the Gaussian component). Indeed the Gabor function gives optimal cover for image area defined by spatial frequency and space axes (Gabor 1946) but the responses of cells in area V1 are not fully described either by difference of Gaussian (DOG) or Gabor filters (Hawken and Parker 1987; Stark and Wilson 1990). In summary, the neural responses of V1, both in the classical receptive field and their modulation by surround stimuli such that changes in texture are signalled more strongly than the texture itself, suggest that they are involved in extracting contour information in the image.

The next cortical areas in the form processing hierarchy of the macaque are V2 and V3, situated in the lunate and inferior occipital sulci (see Figure 1.1). These areas (like V1) contain orientation selective simple and complex cells, some of which are end-stopped, and show clear retinotopic mapping. In V2 and V3 there are more complex (i.e. phase independent) cells. The cells also tend to have larger receptive fields for a given eccentricity than cells in V1 and nearly all are binocular (Zeki 1978a,b; Van Essen and Zeki 1978). A significant number of cells in V2 are also sensitive to subjective or 'illusory' contours defined by co-linearity (44% of tested cells) and continuity (32% of tested cells) (Peterhans and von der Heydt 1989a,b; Peterhans et al. 1986; von der Heydt and Peterhans 1988, 1989; von der Heydt et al. 1992; Peterhans and von der Heydt 1993; see also Fiorani et al. 1992). Recent work suggests that some cells in V1 of both monkey and man may also be sensitive to illusory contours (Grosof et al. 1993). The

responses of these V2 and V3 cells to such illusory contour stimuli maintain the same orientation selectivity as that found using traditional bar and edge stimuli, even when the illusory contour is comprised only of elements orientated 90 degrees from the preferred orientation (Peterhans et al. 1986 and see Figure 1.6, V2 box). That is to say many V2 or V3 cells with a strong response to a vertical line would also respond to a vertical illusory contour defined only by the presence of horizontal lines (see Figure 1.6, box 2). These results suggest that in area V2/V3, a more complete extraction of contour is performed than is seen in area V1.

As in V1, the arrangement of cell selectivities is not random in V2. The retinotopic map can, for each area of visual space, be divided and shown to contain spatially distinct groups of cells responding selectively to motion defined form, static form and colour (Peterhans and von der Heydt 1993). It is of interest to note that this functional division within V2 for form, motion and colour processing coincides with a metabolic biochemical division revealed by staining with cytochrome oxidase in the anaesthetized macaque ('thick stripe' regions containing cells selective for motion, 'thin stripe' regions containing colour selective neurons, and 'interstripe' or 'pale stripe' regions containing orientation selective cells, Hubel and Livingstone 1987; Shipp and Zeki 1985). Recent evidence suggests that this division might not be as clear as first thought in the alert behaving monkey, with orientation selectivity frequently found in thick and pale stripes and also in the thin stripes (Peterhans and von der Heydt 1993). Static illusory contour sensitivity was not found in the thin stripes but was found in the thick and pale stripes. Cell sensitivity for contours defined by coherent motion was found mostly in the thick stripes and some in the pale stripes (Peterhans and von der Heydt 1993). Interestingly, direction selectivity in area V2 of the awake monkey was observed equally but rarely in all cytochrome oxidase defined stripe types of V2, suggesting that V2 plays little role in processing direction of motion (Peterhans and von der Heydt 1993). In summary, although some debate

continues as to the degree of functional modularization within cytochrome oxidase defined regions, the evidence suggests that even within these early visual processing areas there is some division of form and motion processing.

The physiological properties of cells in V1, V2 and V3 suggest strongly that all three areas are involved in the extraction of contours from the image. The cell selectivities seen in V2 and V3 suggest that these cortical areas extract contour information that is not simply defined by luminance. The study of Peterhans and von der Heydt (1993; see also Zeki 1978) provides evidence that the function of V2 does not include processing of direction of motion as previously suggested (Shipp and Zeki, 1985) but rather is confined more to form processing (using motion to define contour information).

Building the building blocks - object-feature detectors

As outlined above, the early cortical visual areas in the macaque brain contain cell populations which are selective for stimulus motion and others selective for stimulus shape (or at least contour information). The anatomical distinction between these pathways becomes much clearer in the subsequent visually responsive brain areas, though it is also clear that the two separate functional sub-systems are interconnected (Young 1992; Martin 1992). Areas MT, MST, FST and STPp all contain neurons which are directionally selective for moving stimuli and will not be considered further in this section, nor will the parietal lobe areas 7a, 7b, LIP and VIP to which they project.

In tracing the properties of from processing further in the ventral stream it is necessary to consider the properties of area V4. Colour, like texture, forms an important part of form processing since it can be combined with shape information to facilitate object recognition (M. Goodale and K. Humphrey, personal communication 1993; E. Brodie, personal communication 1993). Cortical area V4 contains many neurons that are selective for stimulus colour. Zeki (1980) has argued that, unlike the neurons in V1, V2 and V3 which are

sensitive to wavelength, most colour selective neurons in V4 show colour constancy. That is to say, the neural responses in area V4 appear to be independent of the colour (wavelength composition) of the ambient light but rather code the reflectance of different wavelengths of one stimulus region relative to the reflectance of the same wavelengths from other neighbouring stimulus regions in the image. This relative spectral reflectance remains constant independent of the composition of the ambient light, as does the perception of the stimulus colour (i.e. a red object always reflects more long wave light than a green object). The colour constancy exhibited by neurons in V4 indicates interactions with colour coding of areas adjacent to the area inside the cell's classical receptive field (i.e. a surround effect). The property of colour constancy has led Zeki (1980) to call V4 of the macaque the colour coding area (although see Dow 1992), corresponding to the colour centre in man (Lueck et al. 1989).

As shown in Figure 1.3, area V4 acts as a gateway to the temporal lobe areas, which are themselves associated with form processing (see below). Given that inputs to these areas pass through V4, it is not surprising that some cells in V4 are selective for the stimulus form independent of colour while others cells show conjoint sensitivity to both form and colour (Kobatake and Tanaka 1992, 1993; Zeki 1978; Van Essen and Zeki 1978; Gallant et al. 1993). In V4 the selectivity for stimulus form goes beyond the simple orientation selectivity seen in V1 V2 and V3. It contains cells which show selectivity for shapes of (slightly) greater complexity (Tanaka et al. 1991, 1993). Like V1, V2 and V3, there is a retinotopic mapping of cells in V4, although the precision of this mapping is not as strict as in the earlier visual areas (Van Essen and Zeki 1978). This breaking down of clear retinotopic mapping could be due in part to the increase in size of the receptive fields. As in V2 and V3, in V4 there is evidence of clumping of colour-sensitive neurons and clumping of form-selective neurons (Tanaka et al. in press) although this clumping may not be as tight as in V2 (Van Essen 1993).

The next cortical areas in the ventral pathway lie in the inferotemporal cortex. Most investigations of the cell selectivity along the length of the inferotemporal cortex have concentrated on the dorso-lateral sections (areas PITd, CITd and AITd of Felleman and Van Essen 1991). Initial reports indicated that in PIT and CIT, like V4, colour is also coded (Zeki 1977). More recent evidence, however, suggests that IT codes conjointly for both form and colour. The only systematic study of the changing properties along the length of the inferotemporal cortex was performed by Tanaka and colleagues (Tanaka et al. 1991; Kobatake and Tanaka 1992, 1993; Tanaka 1993). During these studies, recordings were made from areas V4 and the inferotemporal areas PITd, CITd and AITd. The results showed a clear trend of receptive field size increase from posterior to anterior areas (average of receptive field areas, V4 = 4 degree^2, PIT = 16 degree^2 and a large jump to an average of 150 degree^2 in AIT). Accompanying this increase in receptive field size was also an increase in the required level of stimulus complexity. In V4 and PIT most (70%) of the cells were selective for single characteristics, such as stimulus size, orientation in the picture plane and colour. Note that the pattern selectivity was not necessarily specific for a single oriented element but could be specific for a line intersection (e.g. 'X' but not '/', '\', '+' or oriented T junctions). 'Elaborate' cells in CIT/AIT showed selectivity for stimuli having greater complexity than seen for cells in V4 and PIT (Tanaka et al. 1991). The simplest of these were selective for patterns such as star shaped stimuli (similar in appearance to many of the stimuli described by Schwartz et al. 1983; Gross 1992). Perhaps the more interesting Elaborate cells in the study of Tanaka et al. (1991) were those that showed selectivity to combinations of features, such as a brown circular region containing a stippled pattern combined with a bar extending to the right; or for a different cell, a dark circular area above a slightly larger white circular region. The frequency of these Elaborate cell types showed a clear relationship with location of the recordings: 2% of V4 cells, 9% of PIT cells and 45% of AIT cells were classified as Elaborate. Supporting evidence that IT

cells are selective for shape and not other stimulus parameters comes from Sary et al. (1993) who showed that for many cells in IT, the same selectivity was present for a neuron whether the shape (e.g. a star) was defined using luminance, motion or texture cues. This indicates that IT cells respond to the shape of the stimulus, regardless of the visual cues used to define the shape boundary. While this obviously shows convergence of inputs defining contours using different visual cues, it also indicates a more abstract and unified coding of 'shape' than seen in the earlier cortical areas. The selectivities of the IT cells can be likened to the "geons" described by Biederman (1987). It is note worthy that the selectivities of IT cells show greater variation and range of stimulus attributes than suggested by Biederman. Further more, the cells described by Tanaka et al. (1991) showed orientation selectivity, whereas Biederman proposed orientation insensitive representations of "geons".

As observed in other visual cortical areas, there was also evidence from the study of Tanaka et al. (1991) of a clumped, modular or columnar organization in AIT: during penetrations perpendicular to the cortical surface, cells were found to have similar selectivities, whereas 2 or 3 mm away (along the cortical surface) a second penetration would reveal cells with substantially different stimulus selectivities. This led to the suggestion of shape processing modules within AIT, with many different modules coding the overall image contents, and interactions within these modules coding precise details of each element (Fujita et al. 1992; Tanaka et al. 1993; Tanaka 1993; Young 1993; see also Perrett et al. 1984; Gawne and Richmond 1993).

The studies by Tanaka and colleagues (Tanaka et al. 1991, 1993; Fujita et al. 1992; Kobatake and Tanaka 1992a,b, 1993; Tanaka 1993) demonstrate several important points. The stimulus selectivity becomes progressively higher moving from posterior to anterior inferotemporal areas. The increase in selectivity can be seen as a sequence of relatively small steps (e.g. from single oriented elements, to T junctions, to more complex star like shapes to combinations of shapes). Each of

these steps could perhaps be accomplished by combining the output from several cells at earlier areas with different simple selectivities. The additional colour or texture sensitivity seen for some inferotemporal cells could also be accomplished in a similar way, combining the outputs of cells with colour or texture sensitivity with the outputs of cells selective for shape.

It has been reported that there is a log-polar transformation of the visual field onto V1 (Schwartz 1980). From an engineering perspective this transformation helps solve size and orientation constancy (Seibert and Waxman 1991, 1992a,b). If the centre of mass of an object is fixated and the image undergoes a log-polar transform then changes in image size and orientation translate into shifts along the 'log' and 'angle' axes. Given the tendency for parallel projections from one cortical area to the next and the limited spread of the axonal arbourizations it is relatively simple to form associations between activities in adjacent cortical areas (e.g. a translation along the 'log axis' or 'angle axis'). Such associations would permit generalization of arbitrary patterns over size or orientation. With the log-polar transform, however, the problems associated with extracting positional invariance for stimulus locations outside the foveal area remain.

Interestingly, orientation constancy (in the picture plane) and size constancy are not generally present in the responses of IT Elaborate cells. Only 3% showed orientation constancy responding equally to presentation of effective stimuli at all angles. A similarly small number showed constancy over a four-fold size change, while only one cell showed size constancy over an eight-fold range of size change (Tanaka et al. 1991). A further notable feature of the Elaborate cells of Tanaka (1991) is that the stimulus selectivity is present over the whole receptive field (i.e. generalization over stimulus position on the retina). Since moving the components of the effective stimuli relative to each other reduced response magnitudes, pattern configuration selectivity is implied over the whole receptive field. It should be noted that other investigations have shown that there

is slight modulation of response magnitude (but not response selectivity) within the receptive fields of inferotemporal neurons (Gross 1992). The results from these studies suggest that the visual system may have log-polar like transformations for another reason (e.g. Baddely 1993). From properties of IT cells it seems that the biological system establishes positional invariance before size and orientation generalization is developed, exactly the reverse of what might be expected from the potential advantages of a log-polar transform.

The properties of the Elaborate cells in the inferotemporal cortex (particularly AIT) suggest that they could form the basic building blocks of object recognition (Perrett and Oram 1993). Tanaka and colleagues (Fujita et al. 1993; Tanaka et al. 1993) argue that the activity across different modules could code the general 'class' of an object feature, whereas within a column of Elaborate cells the small differences in neural selectivity would enable precise signalling of the exact nature of the visual features present in an image. The basic idea of within and between column/module coding is illustrated in Figure 1.4. In the upper part of the figure, four columns are represented, each containing four cells with slightly different selectivities. As can be seen, the differences of the selectivities of cells within any given column is smaller than the differences of selectivities between columns (i.e. the columns are functionally differentiated). In the lower section of Figure 1.4 are two examples of 'images' or object instances that can be distinguished using just five of the 'feature detectors' in the upper part of the figure.

The scheme shown in Figure 1.4 is, of course, greatly simplified: in particular there is very little configurational information. There are two points to be made here. First, the addition of modules which contain cells with configurational selectivity could greatly restrict the number of allowable configurations of the input feature set. For example, Tanaka (1991) found several examples of cells which showed selectivity for concentric 'eye like' patterns. Input from cells with this type of response characteristic would constrain the likely

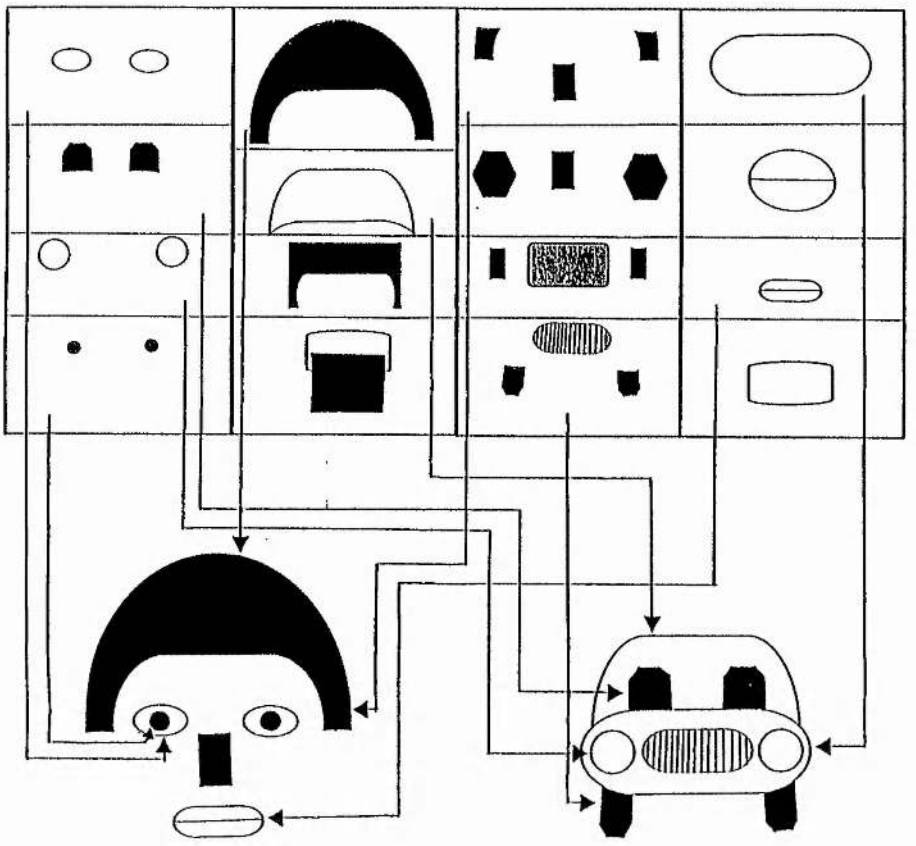


FIGURE 1.4. CODING OF OBJECT-FEATURE AND OBJECT CLASS IN INFERO-TEMPORAL CORTEX. UPPER: COLUMNS OF OBJECT-FEATURE DETECTORS. Four columns are schematically illustrated, each with 4 cells showing slightly different selectivties for object-features. LOWER: TWO POSSIBLE OBJECT CLASSES. Two simplified versions of preferred stimuli are shown. The arrows from the object-feature detectors in the upper part of the figure indicate how two different objects (face and front view of a car) might be coded by combining the output from the object-feature detectors.

relationship of the inputs from the left-hand column to the face shape in Figure 1.4. In other words a detailed configuration specification could be constructed from several inputs each containing some, possibly imprecise, configuration information (e.g Fukushima 1980). Also the number of objects experienced which would share all of a given set of features is small. In other words the visual system may not have to operate with perfect configuration sensitivity: the system can 'guess' that if a set of 20 object-features are present, then a particular object (e.g. a front view of a car) is very likely to be present.

It should be noted that recent findings of low levels of correlation of noise between neighbouring neurons suggests that only a small number (maximum of 20) of IT cells would be able to contribute to the signalling of the precise shape of any one feature (Gawne and Richmond 1993). It is likely therefore that within any particular module only a small number of cells can encode the feature presently being viewed because of this correlated noise. Tanaka and colleagues estimated that there would be at least 600 of these modules in macaque IT cortex. Even allowing for simple binary coding using this visual shape alphabet with one feature coded per module, this gives a staggering number of possible pattern combinations (2^{600}, which is more than the estimated number of atoms in the universe!)

Grandmother cells and the coding of biologically important objects

There is a further class of cell found both in AIT and STPa which shows selectivity for biologically important visual patterns. These cells are selectively responsive to the head, face, hands and limbs (e.g. Bruce et al. 1981; Baylis et al. 1985, 1987; Gross et al. 1986; Desimone et al. 1984; Rolls 1984; Perrett et al. 1982, 1985, 1989, 1991; Yamane et al. 1988; Hasselmo et al. 1989; Young and Yamane 1992). While some of these cells are selective for particular features of the head and face (e.g. the eyes, Perrett et al. 1982, 1992), others are selective for multiple parts of the same object (Perrett et al. 1984, 1993, Wachsmuth et al.

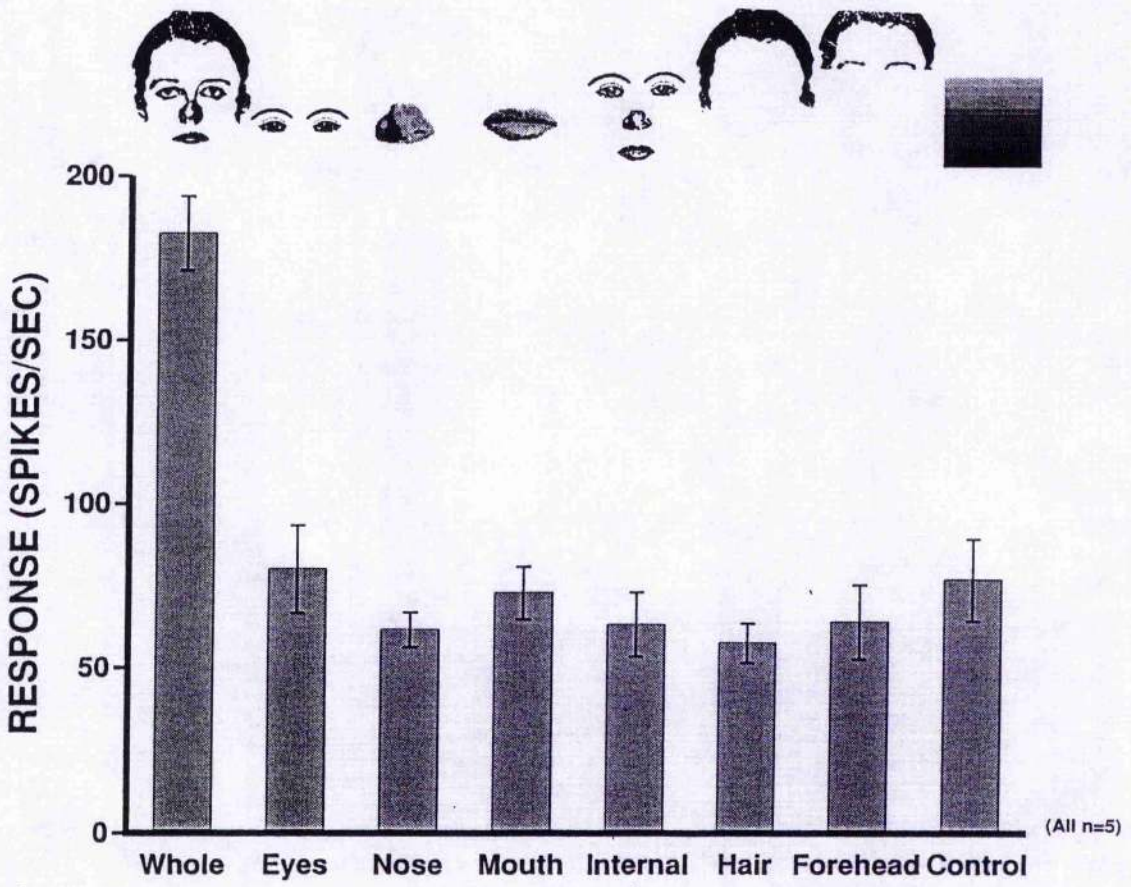


FIGURE 1.5. AN EXAMPLE OF A CELL SELECTIVE FOR FACE STIMULI. Mean response magnitudes (+/- S.E.M., n=5) of one area STPa cell to a face and its components. The stimuli are represented across the upper section. As can be seen, the cell produces a strong response (>175 spikes/second) to the whole face, yet produces no response above control level to the component features when tested in isolation.

1993; see also Figure 1.5). The cells selective for particular facial features could perhaps be viewed as an instance of Tanaka's Elaborate cells but, certainly in some cases, the cells have other selectivities which suggest that they are coding more complex and abstract information (e.g. the direction of another's attention, combining information about eye gaze angle, head direction and body posture all within the response of one cell, Perrett et al. 1992).

The cells responsive to faces located within AIT show selectivity for orientation (e.g. one cell may respond more to full face images than to images of either the left or right profile) and are sensitive to rotation in the picture plane, with a strong bias for selectivity for upright faces (Tanaka et al. 1991). Further, as with the Elaborate cells of AIT they do not show generalization across large changes in size. The results from studies of face selective cells in STPa suggest that, in contrast, these cells show much greater generalization for both retinal size and orientation of the head in the picture plane (Perrett et al. 1982, 1984, 1988; Rolls et al. 1986). Cells in IT and STPa therefore show different levels of generalization (see Figure 1.6). The STPa cells responsive to faces also show constant tuning across different spatial frequency bands (Rolls et al. 1985, 1987). Studies of the effects of the direction of the ambient light have shown that cells of STPa show constancy across different directions of ambient light, even though the visual appearance of the head changes dramatically with these different lighting conditions (Hietanen et al. 1992). The lighting constancy exhibited by STPa cells might not be as surprising as it first appears, since it may depend on luminance and contrast normalization for local regions in an image which could occur during early visual processing in the retina, LGN and V1 (e.g. Hubel and Livingstone 1990). These processes would facilitate simultaneous contour extraction across the whole image, independent of marked regional variations in light level, thus enabling the same contour and feature processing inputs to be active regardless of lighting condition.

Cells found in either area AIT or area STPa do not tend to show complete constancy across change in perspective view. Furthermore different cells are maximally tuned to different head views (Desimone et al. 1984; Perrett et al. 1985, 1991; Hasselmo et al. 1989a). Interestingly, a preponderance of cells are tuned to one of only 4 views (full face, left or right profile, back of the head), suggesting that view tuning is not randomly distributed but rather that certain 'characteristic views' are preferentially coded (Perrett et al. 1991). There are, however, exceptions to this sensitivity to perspective view in area STPa: a small percentage of cells studied (< 5%) have been found which respond to any view of the head but not to control objects matched for size, complexity and capacity to arouse (Perrett et al. 1984, 1991; Hasselmo et al. 1989a). These cells display selectivity characteristic of object-centered representations of one object part. Sensitivity to differences between individual faces (i.e. identity) in areas AIT and STPa may be coded across populations of cells and is rarely found at the single cell level, particularly in area STPa (but see Young and Yamane 1992a,b; Baylis et al. 1985; Hasselmo et al. 1989b; Rolls et al. 1989; Perrett et al. 1984, 1989). Note that the small percentages of the cells studied do not reflect the possible importance of the population's function. Indeed cells showing object-centered, identity-specific selectivity (i.e. grandmother or cardinal cells) would be expected to be found only rarely (e.g a total active population of only some 1000, Barlow 1972). However, the very existence of a much greater number of cells with view-specific, identity-general selectivities suggests that this stage in the processing of visual information perhaps subserves a greater number of processes (e.g. social signals and integration with motion inputs as well as object recognition, Perrett et al. 1993; see chapter 7).

Tanaka et al. (1991) reported cells in the inferotemporal cortex (IT) which required specific combinations of features, not only to be present but also to be in particular spatial relationships. Recently, Miyashita et al. (1993a) has reported cell selectivity within ventral IT indicating that the effect of stimulus

rotation/reflection seems to have less of an effect on response magnitudes than changing the relative spatial arrangement of the component features. With spatial selectivity as well as simple feature content, output from such IT cells could be used to establish the configuration selectivity for head view seen in area STPa (e.g. Perrett et al. 1984). A direct demonstration of selectivity to configuration of STPa cells sensitive to the face has been shown by Perrett et al. (1982), who found that many cells were less responsive to faces when the features (eyes, nose, mouth) were jumbled within the outline of the face.

In summary the changes in cell selectivity when moving from area AIT to area STPa seem to be greater size, orientation and possibly also lighting constancy. A change from response sensitivity to perspective view (viewer centred) to responses independent of perspective view (object-centered), however, does not seem to occur except for the possible exception of a small number of cells within area STPa. Generalization across different parts of the body (e.g. head, torso and legs) also appears to be present in area STPa (Perrett et al. 1992; Wachsmuth et al. 1993).

Rationale for multistage form processing

It is of interest to speculate about the 'jumps' that appear to be available in the processing scheme from anatomical considerations (see Figure 1.3). Note that while it is possible to 'miss' V2/V3, there is no anatomical link (and hence no functional link) that allows contour information from V1 and V2 (stage 1) to project to cells with moderately complex form selectivity (e.g. V1/V2 to PIT). Thus V4 seems to mark the beginning of building the building blocks (stage 2). Likewise output from V4 must pass through either PIT or CIT (or both) before reaching AIT. This suggests that there is a need for at least one area in the processing of the building blocks (in PIT/CIT, stage 2) before cell selectivities can be combined to form selectivity for putative object-feature instances (AIT, stage 3). In the same way generalization across object instances (stage 4) does not

seem possible directly from complex feature detectors, there being no connection from stage 2 (PIT and CIT) to STPa.

The lack of direct projections from the lowest (e.g. V1/V2) to the highest visual areas (e.g. AIT/STPa) also initially seems surprising since convergence of all supportive information would provide greater evidence of an object's presence. The selectivity of cells in these early visual areas, however, is such that any one cell would be activated by virtually every natural image. Therefore the information that they code is unlikely to aid directly complex object identification. Note that only a small proportion of the total number of cells in V1 would be active for any one image, but that any one cell could easily be active for many images. For example, trees, faces, buildings, cars and chairs all contain vertical edges. The information about the presence of vertical edges tells you a lot about edge orientation but very little about the identity of the object. Therefore to build a processing system capable of extracting object identification using only edge orientation information would require many contingencies to be accounted for. One contingency might be "if there is a horizontal edge at retinal location (X,Y) then there should also be a vertical edge at (X+2,Y)". Another related contingency might be "there is not a vertical edge at (X+1,Y) and not at (X,Y+1) and not at.....". In other words cells at the higher levels representing objects would have numerous 'and' plus 'and not' requirements. This amounts to the familiar 'combinatorial' or 'convergence explosion'. Further, such a system would need to embody a complete compliment of contingency processing sets for all orientations, sizes, positions, lighting conditions and so on. While individual neurons can receive many inputs (1000-10,000) the number is far too small to accommodate the total number of inputs that would be required to satisfy all contingencies in terms of simple edges.

The four stage, multi-layer model outlined above avoids this 'convergence' or 'combinatorial explosion'. By gradually developing stimulus selectivities, the number of units activated by a natural image decreases at higher processing stages

(Barlow 1985). The combinatorial explosion problem can be likened to the game of 20 questions. If one is allowed 20 questions in order to ascertain what object is present in an image, it would not be sensible to ask about the presence of small oriented edges at particular positions. One would need to ask a huge number of questions of this type. On the other hand, one could derive knowledge about simple shapes (e.g. 'H' 'T' 'Y' and '+') within restricted locations from 20 questions about local line orientation. If one then asks 20 questions using this knowledge, one is in effect asking 400 questions about oriented edges. At the next level the number effectively increases to 8000 and so on. After 6 steps in processing (with just 20 questions per step) each STPa cell would be effectively asking 64,000,000 questions about local edge orientations over a wide area on the retinal.

Thus a multi-stage system allows for relatively simple connection patterns between processing elements to be used to establish feature detectors of ever increasing complexity. It would be of interest to discover what combinations of simple orientation features provide the best discriminatory power between natural images: predictions based on the above arguments suggest that they will be similar to those found in areas V4 and PIT, including 'T' and 'Y' junctions and crosses.

Learning associations and generalizations

An important series of studies into the mechanisms by which learning affects the selectivity of cells of AIT cortex was performed by Miyashita and colleagues (Miyashita and Chang 1988; Sakai and Miyashita 1991; Miyashita 1990; Miyashita et al. 1993b). During an extensive training period in which 97 visually distinct coloured abstract (fractal) patterns were learnt, the monkeys had to indicate whether a particular pattern was the same or different as a sample presented 16 seconds earlier. During this training, the 97 sample fractals were presented in the same order. Recordings from temporal cortex after this training indicated that some cells were selectively activated by particular learnt fractal

patterns in the stimulus set. The choice of fractal patterns eliminated any particular biological importance of the patterns except for the associated behavioural implication and thus indicated effects of experience on the shape selectivity of single cells. Of particular interest here was the observation that cells were selective not for one but for two or three of these fractal patterns. The response selectivity was for patterns that were adjacent in the sequence of presentation (e.g. patterns 91-93 or 56-59, but not 56 and 91), and therefore this learnt selectivity could not be explained by similar visual appearance (Miyashita and Chang 1988; Sakai and Miyashita 1991; Miyashita 1990). Other studies indicated that fractal patterns arbitrary paired could also come to activate the same AIT cells (Sakai and Miyashita 1991) provided that the pairings were behaviourally relevant (Miyashita et al. 1993b). These studies are important, since they show not only plasticity in ventral stream neural responses but also that associative learning can occur at the single cell level between substantially different visual images.

A learning scheme using 'trace' activity levels to establish positional invariance simultaneously for each of many experienced features has been proposed by Foldiak (1991). The trace activity allowed association of one feature detector with a second detector that became active soon after the first. In the example given by Foldiak (1991), four differently oriented line detectors were present at each of 8 by 8 'retinal' locations in the input layer. During learning, these were activated by long oriented lines sweeping in random directions across the input array. After learning, the outputs showed large receptive fields formed by Hebbian association between input and output layers. Importantly the selectivity of a given output unit was for only one of the four possible oriented lines seen in the input detectors. This can be achieved by decorrelating inhibitory connections between the output units using a form of anti-Hebbian learning (Foldiak 1989, 1990).

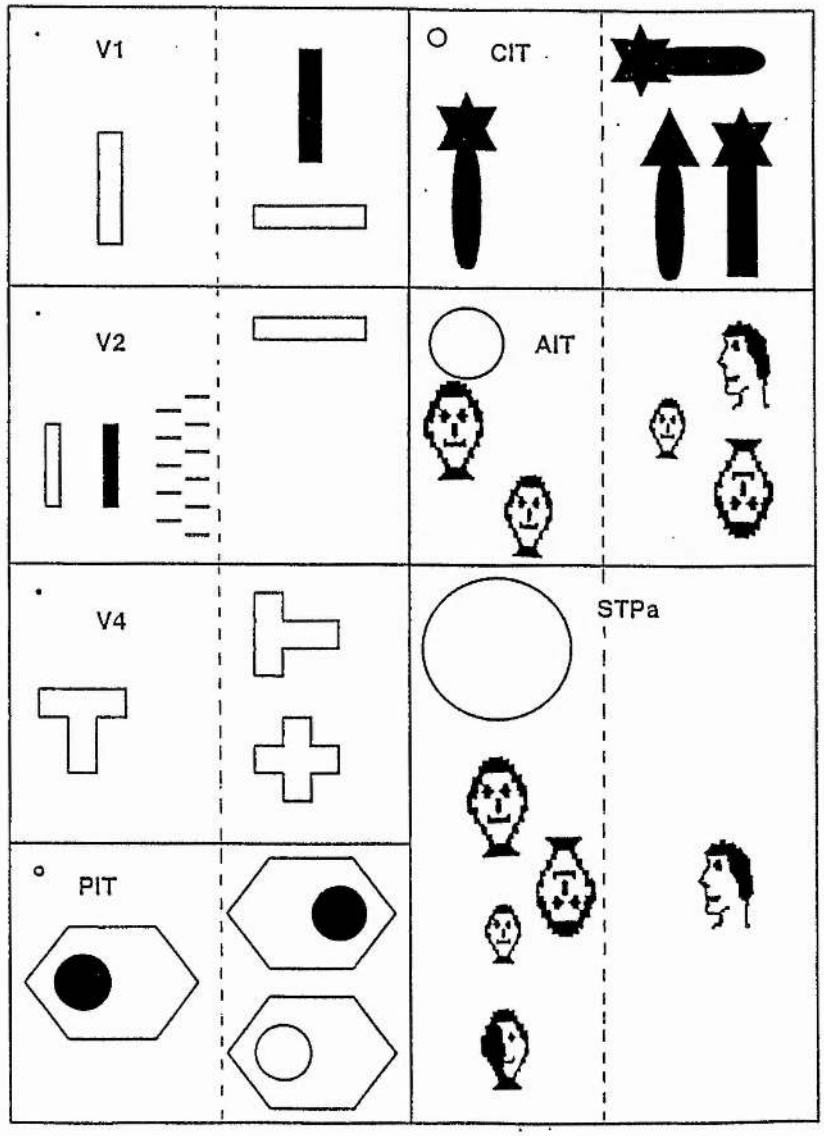


FIGURE 1.6. SUMMARY OF NEURAL RESPONSE SELECTIVITIES IN THE FORM PATHWAY. Idealised single unit response selectivities for areas in the form pathway are shown. The left section of each box is an example of an effective stimulus. The right section shows examples of stimuli that would be ineffective. The circle in the top left represents the relative receptive field size. See text for details.

A mechanism combining Hebbian and anti-Hebbian mechanisms could learn generalization across perspective view of several objects from experience of rotation of these objects, given inputs with selectivity to different views of each object. Hebbian association, utilizing trace activity, would establish selectivity for any view of an object, while the decorrelation between outputs would ensure that each output node was sensitive to only one object. In principle, such a scheme can also learn other invariances, such as lighting, size and orientation providing that the appropriate input detectors exist. It is therefore possible that a simple unsupervised learning scheme, such as Hebbian learning combined with decorrelation (Foldiak 1989, 1990; Barlow and Foldiak 1989) could be used to develop the proposed transitions leading to the progressive generalizations from IT cortex to area STPa.

That the existence of modules or columns within IT might arise from decorrelation of inputs. Indeed, it is conceivable that within any given module, the individual cell selectivities are themselves decorrelated, thereby establishing a sparse representation of the possible object-feature instances that comprise an object-feature class (see Figure 1.4 upper). An attractive property of Foldiak's learning scheme is that it operates on the statistics of the input images in a way that is not dissimilar to networks performing principal component analysis (Foldiak 1989). Therefore such a scheme, operating at any level within the visual cortex will extract features that form sparse representations which maximally differentiate images. It will be of great interest to see if such a scheme operating in a multi-layer network can extract features similar to those seen in V4 and subsequent inferotemporal areas.

Summary of form processing in the ventral stream

Figure 1.6 gives a simplified summary of the neurophysiological properties of cells in the ventral or 'form processing' stream, where effective stimuli for one example cell are shown to the left of the dotted line in each box

and non-effective stimuli are displayed to the right. These non-effective stimuli would, however, be effective for other cells in the same brain area. Computationally the sequence can be viewed as following stages: first extract edges and contours (V1), then search for more general contours (V2) to define simple edge configurations (V4). These can then be combined first to give simple shape selectivities (PIT), which can themselves be combined to produce sensitivity to more complex shapes (CIT). A further combination could give rise to selectivity for very complex abstract-forms and approximate object-feature instances (in area AIT) which would be specific for orientation, size, lighting and perspective view. Combining outputs of many approximate descriptions of these abstract-form and object-feature instances could establish descriptions generalizing across orientation, size and lighting in a view specific manner. A final stage combining different views and object parts (within area STPa and analogous areas) would obtain a viewpoint invariant representation of the object (see Perrett and Oram 1993 for a detailed discussion). In summary the physiological evidence suggests 4 computational stages: extracting contour information, building a 'shape dictionary', generating descriptions of approximate object-feature instances and establishing object representation in a view specific manner. Subsequent stages may generalize across view and infer complete object identity from the presence of object parts. The transitions between all of these proposed stages involve pooling related inputs from the preceding stage such that increased stimulus complexity is generated with simultaneously increased generalization over retinal position, retinal size and lighting conditions. Association of images over time is a possible candidate for a learning system that would explain the development of generalizations seen in the higher stages of processing.

The effects of expectation

When an expectation is met, the phase "I told you so" is often used (especially when something unpleasant happens to someone else). It is often found that within psychological models of recognition top-down influences are postulated to influence visual processing (e.g. Lowe 1987). This section reviews the role of expectation as one form of top-down processing and some of the physiological work that is related to this phenomenon in which the responses of cells change with predictability of the stimulation.

For cells with simple stimulus selectivities, there are data from areas V1, V3 and area 7a of the parietal lobe of relevance to this issue. The sensitivity of V1 cells to surround effects (Allman et al. 1985; Knierim and Van Essen 1992) can be viewed as a form of expectation, since in many images it would be expected that texture in many image areas would be similar (except at object or object-part boundaries). The suppression of responses by elements of a textured surround having similar orientation to the element falling in the classic receptive field can therefore be regarded as response suppression to expected or predicted stimuli.

Galletti and colleagues (Galletti et al. 1990) found that nearly one half of the directionally selective cells in area V3a showed differential responses to expected and unexpected stimulation. Cells in V3a were sensitive to the direction of motion when tested with an oriented bar moving across the visual field and the monkey maintaining steady fixation. The same retinal stimulus motion could also be achieved by having the monkey move its eyes across a static stimulus. For 48% of the cells the response magnitude was reduced under the circumstances when the monkey induced the retinal motion. In area V3a therefore it appears that if image motion is a result of self-produced eye movements (i.e. expected), then the responses of the cells are reduced but, when the stimulus is not expected (i.e. occurs independently of the monkey's behaviour), there are strong responses.

Work of Goldberg and colleagues on cells of the parietal cortex (area 7a) also show a pronounced effect of 'expectation'. Duhamel et al. (1992) recorded from cells within the parietal lobe that showed selectivity for stationary bar and

dot stimuli. In one experimental condition, the stimulus was shown to the monkey but outside the recorded cell's receptive field. The monkey then had to make a sudden, fast eye movement (a saccade) to a new position, such that the stimulus was now, suddenly, in the cell's receptive field. For some cells a surprising result was observed: the cell would start responding before the stimulus fell within the receptive field. This effect could also be observed even when the stimulus was turned off just as the monkey began the saccade but not when an equivalent saccade would result in the receptive field remaining empty. This phenomenon can also be viewed as reflecting expectation, in this case the expectation of visual stimulation in part of the receptive field resulting in an earlier response.

The phenomenon of expectation influencing cell responses is not restricted to the early visual and parietal cortex. A population of cells in area STPa is responsive to the motion of an object but shows no selectivity for the shape or size of the object (Bruce et al. 1981; Perrett et al. 1985; Hikosaka et al. 1988; Mistlin and Perrett 1990; Hietanen and Perrett 1992; Oram et al. 1993; see chapter 5). Such cells show responses to any object moving, except under circumstances when the object motion can be interpreted as 'expected'. For many cells, the response to any object moving up into the line of sight would produce a large response. There was a situation where this response was not seen, namely when the monkey raised its own hand into sight (Hietanen and Perrett 1993). The lack of response was not simply a general suppression of all visual responses, since if an object was brought into view simultaneously with the monkey's hand then the cell would still respond (Hietanen and Perrett 1993). This indicates that the expectation effect is tied not only to the motion but also to the form and location of the moving stimulus. Similarly, when the monkey controlled the motion of a visual pattern by rotation of a handle held out of sight, responses for some STPa cells did not occur, whereas the same stimulus motion occurring when the experimenter rotated the handle produced strong responses (Hietanen 1993; Hietanen and Perrett in prep.). These results indicate that, for some STPa cells,

any stimulus can produce a response when its motion is unpredictable but the same stimulus fails to produce a response if its motion is predictable from the monkey's own motor behaviour.

Area STPa contains cells that are selective for tactile stimuli as well as visual stimuli (Bruce et al. 1981; Hikosaka et al. 1988; Mistlin and Perrett 1990). The receptive fields of these tactile sensitive cells are typically very large and can include the whole body (Hikosaka et al. 1988). The responses of these cells, however, can be modified by visual inputs: the response to tactile stimulation from an object touching the subject after the object is seen approaching the area to be stimulated is greatly reduced compared to the response when the same stimulation occurs out of sight (Mistlin and Perrett 1990). In this situation, the difference between the two conditions is only the sight or lack of sight of the object's approach before tactile sensation (expected or unexpected stimulation). Furthermore, vision was found to be only one modality providing a source of expectations about impending tactile stimuli. Normally the monkey was seated on a smooth chair during experimentation and was familiar with the tactile nature of the chair. Cells showing tactile responses would not respond as the monkey touched different parts of the chair, nor would they respond when the monkey touched its own fur. These cells did respond when a small part of the chair was covered (out of sight) with a new material (e.g. fur) and the monkey touched this new 'unexpected' texture (Mistlin and Perrett 1990). In these situations, predictions as to the likely tactile qualities of different parts of the environment were based on memory. These predictions were effectively negating high level sensory processing whenever predictable tactile experiences were encountered.

The data on 'expectation' and its influence on cell response properties therefore suggest that expectation has powerful effects on STPa cell responses. These effects are not restricted to visual processing, since visual cues can also affect tactile processing. Such effects of 'expectation' on motion sensitivity also seem to be present in visual areas V1 and V2 but are more prominent in later

areas (e.g. V3a, MT and MST, Galletti et al. 1984, 1988, 1990; Erickson and Thier 1991). The effect of expectation on static information processing do not, however, appear in the earliest visual areas (R. von der Heydt, personal communication, August 1993). For STPa cells in the ventral stream, only the unexpected or unpredictable stimuli produce responses. It is of interest that area 7a in the dorsal pathway shows a very different effect where expected stimuli provoke a response 'in advance' of their actual occurrence: further work needs to be performed to see whether expectation effects in both dorsal and ventral processing streams have some common underlying mechanism.

Coding distinctive attributes relative to the norm

Consideration of the temporal constraints suggest that the visual system can operate in a purely feed-forward way during processing of complex patterns which are suddenly presented. As already pointed out, anatomically nearly all connections between cortical areas are reciprocal. Indeed there are as many if not more feed-back connections from area V1 to the lateral geniculate nucleus (LGN) than there are connections from the LGN to V1. Surprisingly, there is very little known about the function of these feed-back connections. The temporal constraints outlined above suggest that the function of feed-back connections may have little to do with visual processing under the circumstances of suddenly presented objects. Therefore one possible role of feed-back can be viewed as eliminating the "I told you so" element of processing, allowing only the unaccounted or unexpected information to be processed. Realizations or expectations as to the nature of forthcoming stimulation can come from a great variety of sources. One source of predictable visual images is as a consequence of the animals' own behaviour, as reviewed above. Predictability could also arise from ongoing processing which 'resolves' some of the input characteristics. These resolutions could then steer subsequent analysis. It is possible that effects of predictability from this latter source may partly explain the decay in firing rate

seen in STPa cell responses after an initial burst of activity (Oram and Perrett 1992; Oram et al. 1993; see chapters 3 and 5).

Below one possible role of feed-back is considered by summarizing some psychological data on the efficiency of object recognition in humans. Evidence from psychological studies using caricatures of images suggests that the caricature image might be more recognisable than the true or veridical image (Benson and Perrett 1991; Rhodes et al. 1987). To understand the relevance of this work it is necessary to describe the caricature process briefly. Multiple images of an object class (e.g. faces) can be combined together to produce an average, norm or prototype representation. The caricature process compares the image of an instance within the class (e.g. Fred's face) with a prototype face. Any differences from the prototype can then be exaggerated (e.g. a bigger than average nose can be made even bigger, a smaller mouth made even smaller). The process can be performed on all aspects of the image such as shape, colour and texture either independently or in combination (Rowland & Perrett in prep.).

When subjects are presented with such slightly exaggerated or caricatured images, they are frequently able to recognise the identity of the image more quickly (often accompanied by an increase in accuracy) than when the veridical image is presented. This caricature advantage (as it is called) has been interpreted as support for the notion of a prototype of an object class being stored in memory with individual instances being represented as 'distances' or vectors from that prototype.

Although further work needs to be performed to clarify which aspects of the image give the greatest caricature advantage, the phenomenon suggests a possible style of processing for feed-back connections in image pattern processing: that of removing resolved image features from further processing. For example, Seibert and Waxman (1993) have developed a model which, as a first pass, categorises image classes (e.g. face). The model then uses the difference between the classified input and the stored face 'prototype' to extract the

unresolved or unaccounted image elements. This extracted difference information is then passed through a similar process allowing further within category distinctions (e.g. Fred's or John's face).

In a similar manner, Pece (1993) has proposed that one functional role of the feed-back from V1 to LGN might be to balance the input to the LGN with the processed output from V1. Again this would mean that only information that was not already resolved and passed on would be passed forward during a second phase of processing. Such feed-back of information could of course occur between the output and input layers of one cortical area as well as between cortical areas. Evidence from the cat suggests that feed-back influences from area 18 (V2) to area 17 (V1) tend to have maximal influence on cells with similar orientation preference and receptive field position (Alonso et al. 1993).

To speculate on the benefit of such a process, one can imagine a module or column in IT with many cells each selective for a pair of features arranged horizontally (e.g. Figure 1.4, first column). While many stimuli might activate cells within this column, these stimuli would all have two separate areas arranged horizontally. Activity from the whole column would represent a prototype feature (i.e a horizontal pair of blobs). If the total output from a column or module provides inhibitory feed-back on the column itself, then much of the activity within the column would be negated. The remaining activity would represent the distinctive attributes of the input image *relative* to the feature prototype. Thus the particular type of feature pairing would be accentuated by comparison to horizontal paired features in general. Such norm-based coding is designed to extract the distinctive features of a given pattern relative to the average or norm of the class of patterns to which it belongs. The functionally defined columnar architecture which could implement norm-based coding of input patterns is evident throughout the cerebral cortex. Norm-based coding may therefore be an important general principle in neural computation.

This speculative scheme leads to two predictions for the dynamics of cell responses. There should be a general decline in activity for all cells initially activated by a stimulus and a faster decline of activity for cells responding to sub-optimal stimuli. Both these predictions are evident in the time course of STPa cells (Oram and Perrett 1992; see chapter 3), though of course similar predictions would be made from lateral or feed-back inhibitory connections onto these STPa cells.

As already noted, the number of columns containing object-feature detectors in IT is high. This means that many images could be discriminated on the basis of column activity, without need for the type of processing proposed above. Furthermore, if a decorrelating learning mechanism acted between these columns, then the maximum possible number of images could be distinguished in a feed-forward way. It is worth stressing again that using many columnar object-feature signals in conjunction can show remarkably accurate and rapid discrimination between input patterns (see chapter 3).

At present, of course, the above type of processing remains speculative. For example, attention to other stimuli is also known to cause a reduction in firing rate in V1, V2, V4 and IT (Richmond and Sato 1987; Richmond et al, 1983; Moran and Desimone 1985; Sato 1988, 1989; Chelazzi et al. 1993; Motter 1993), although this alone is unlikely to produce the differential decay in firing rate between different stimuli. Establishing the functional roles of feed-back connections within the brain remains one of the most intriguing and possibly important areas of research: at present they remain something of mystery.

The overview given of the biological data underlying shape processing gives three spheres of constraints for models: (1) anatomical considerations suggest architectural constraints with a minimum of 6 cortical areas before accessing the highest levels, (2) the physiological properties of cells within the

cortical areas associated with form processing suggest a series of computational stages, often involving two or three steps, which should be performed and (3) temporal response properties suggest the types of connection by which the cellular selectivities should be established. Below these are brought together to form the outline of biological processing of static visual patterns.

Four stages of processing

The anatomy of the primate form-processing pathway indicates that the visual system employs a multi-layer network. The precise number of layers involved is debatable, although it is suggested that there are seven or eight major cortical areas, each of which involves an input and output layer. The actual number of layers within each cortical area is potentially greater, although considerations of the speed of processing and response latencies suggest that there cannot be more steps in the first pass of processing.

The connections between the areas of primate visual form processing suggest that possibly only four computational stages are needed: contour extraction and feature grouping at the first stage (V1-V2/V3), development of relatively complex 'feature detectors' at a number of retinal locations during the second stage (V4, PIT, CIT), establishment of instances of approximate patterns or object-features at a particular size and orientation in the third stage (primarily AIT) and subsequent generalization across object instance in a view specific manner (STPa). Further combining of object parts and different views also occurs in area STPa. It is unlikely that all these stages can be achieved in single steps. For instance, to detect illusory contours and general Gestalt groupings it is first necessary to have processed the actual local contours and other elementary features present in the image. Likewise, the development of object-feature detectors may be easier using inputs which show selectivity for moderately complex shapes. The gradual hierarchical combination by-passes the 'convergence explosion' problem.

The final transition that appears to occur in the processing of object patterns is from area AIT to area STPa. Here many object constancies are established (e.g. across orientation, size and ambient lighting). There does, however, seem to be one constancy that is rarely established: constancy of response over angle of view or viewpoint constancy. In the few cases where generalization across perspective view occurs, this may happen after generalization across size, orientation, lighting condition etc., since there is some evidence that the response latencies of object-centered cells in STPa are slightly larger than those of view-specific cells (Perrett et al. 1992).

It is suggested that the third stage of processing should show nodes with sensitivity to what can be called object-feature instances. This can be regarded as selectivity for an input pattern that approximates a real 3-D object or a feature that is normally indicative of only a few objects (e.g. a cell selective for an upright face stimulus or a cell selective for an eye-like stimulus, described by Tanaka et al. 1991). Such cells are sensitive to pattern elements or object-features that tend to occur under particular viewing conditions (e.g. three or four fold changes in retinal size, perspective view +/- 60 degrees, orientation +/- 45 degrees, etc.). In this model it is conceived that many such approximate descriptions exist within temporal cortex and are activated by each object class that is viewed. These object-feature detectors are organised by similarity of the feature rather than similarity of parent object, although for objects of particular importance (e.g. faces) it is likely that such detectors form their own distinct groups or modules.

The scheme above follows object recognition as a bottom-up process. Anatomical considerations reveal architecture for extensive top-down influences in visual processing. We have reviewed effects of expectation that are possible manifestations of top-down influences on high level sensory processing. Such effects may be far more widespread but have not yet been studied extensively. Columnar grouping of cells in temporal cortex suggests a second type of top-down influence whereby image attributes are processed relative to prototypes.

Such processing would enhance the ability of the nervous system to realize the distinctive features of objects to enable precise classification. Despite these speculations about the role of feed-back processes in vision, the dynamics of cell responses indicate that the classification and recognition of objects can be achieved with remarkable accuracy and speed in a purely feed-forward manner.

CHAPTER 2

GENERAL METHODS

TRAINING

The subjects were juvenile Rhesus macaque monkeys (*Macaca mulatta*), weighing 3-8 Kg. Individuals were reared within a social colony at a U.K. registered breeding colony. Experimental animals were first trained to sit quietly in a primate chair and accept food and liquid from the experimenters. Once the subject would enter the primate chair of its own accord, then training on a colour discrimination task was started.

During the task the subject was encouraged to lick tubes placed in front of the mouth. Training of the behavioural task was encouraged by water deprivation (12 hrs) before each recording session. Once the subject freely took liquid from the lick tubes, the task was changed to incorporate a simple colour discrimination. Two liquid rewards were used: either sweet fruit juice or a weak saline solution. The fruit juice reward was paired with a green LED light, whereas the weak saline solution was paired with a red LED. Initially the LEDs were placed close to the subject, but were gradually moved further away. The subject was regarded as being trained when performance on the task reached a high level (> 80% correct) even when the LED was at a distance of 4 m on an otherwise white wall. During latter stages of training a variety of objects were presented simultaneously with the LED.

OPERATION

All procedures were conducted under Home Office Project licence numbers PPL-60/00353 and PPL-60/01453 and Personal licence number PIL-60/02475.

Sterotaxic implant

The implant contained two stainless steel David Kopf wells and two restraining head bar holders. The position, in sterotaxic co-ordinates of the wells typically were 8-15mm anterior of the interaural line, and 12-14 mm lateral of the midline. Once the precise positional measurements had been decided, a to scale diagram (using graph paper marked out with a 1mm grid) was made, indicating where the restraining bars would be positioned. The elements of the implant were then placed in position on a glass sheet and joined using dental acrylic.

Pre-operation procedure

The subject was put into the travelling cage and given a sedating injection of Ketamine (0.5-1.0 ml of Vetelar). Liquid paraffin oil (Vaseline in liquid form) was dropped into the eyes to protect and prevent drying. The head was also shaved then swabbed down with alcohol and tincture of iodine. Atropine (1 ml of 600 micro-gram/ml) was injected to reduce secretions. A general purpose wide spectrum anti-biotic (1 ml ampicilin) was given as a precautionary measure. An intravenous cannula was inserted allowing direct administration of the anaesthetic (Sagatal).

The operation

The monkey was placed in a stereotaxic frame and rested on a diathermy base plate. A breathing counter and rectal thermometer probe were put in position and linked to a heating plate to maintain body temperature. The operation involved one incision, running along the midline of the skull from just above the eye ridges to the back of the crown. The skin and underlying membranes were reflected away from the skull and held away using the haemostats. Any local bleeding was quarterised using a diathermy needle.

Drilling of the skull was done with constant irrigation to keep the bone temperature down. [If the skull tissue rises above 50 degrees centigrade, then it dies.] Two wells were drilled out as circular plugs. The implant was then positioned to allow for the localization of the T bars and screw placement. The stainless steel T bars and screws were inserted into small slots drilled in the skull. Dental acrylic was then placed around the implant to secure it in place. After allowing time for recovery from the operation (2-4 days), the subject was retrained until pre-operative performance was reached. There after the dura underneath the well caps was swabbed clean every 2-3 days.

RECORDING

Behavioural task

Before recording began, the subjects were retrained to discriminate between the red or green colour of an LED light (see Training section above). The LED was typically situated level with the monkey's line of sight on a blank white wall at a distance of 4 m, but could also be placed +/- 15 degrees to the left or right or +/- 10 degrees above or below this central position. Head movement of the subjects in the primate chair was restrained by passing restraining rods through the restraining tubes in the implant. Recording sessions lasted for periods of 2-4 hours. The LED and test visual stimuli were presented from behind a large aperture (6.5 cm diameter) electromechanical shutter (Compur) or an alternative (20 cm square) liquid crystal shutter (Screen Print Technology Ltd.). Both types of shutter had rise times of < 15 ms. When open the shutter allowed the monkey to view only the central 30 (Compur) or 100 (liquid crystal shutter) degrees of visual space. On each trial the shutter was opened under computer control (after a 0.5 s signal tone) to reveal the stimulus and remained open for a period of 1 s. The LED light became visible at the time of shutter opening (stimulus presentation) and was randomly red or green on different trials. The monkey's

task was to lick for a fruit juice reward when the colour of the LED was green and to refrain from licking when the LED was red to avoid delivery of a weak salt solution. The LED colour was changed in a pseudo-random order under computer control. Subjects were deprived of water for periods of up to twenty-four hours before training and recording sessions to motivate task performance.

Although the subjects did not have to fixate the LED throughout the trial period, the monkeys attended to the LED at the beginning of trials in order to lick several times for multiple juice rewards in the 1.0 s trial period. Once they had judged the colour of the LED they were then free to move their eyes. The 2D test stimuli were projected onto the wall on which the LED was located, 3D test stimuli were presented in front, below or to either side of the LED. The monkeys performed the task at a high level of accuracy during the recording sessions independent of simultaneously presented test stimuli. On trials where the monkey licked for fruit juice, normally two and occasionally three licks were completed in the 1 second period available.

Stimulus presentation

Various types of visual stimuli were presented while the monkeys performed the behavioural task (see above). Trials were initiated by the experimenter but thereafter under computer control and consisted of a 0.5 second warning tone, followed by the shutter opening for 1 second to reveal the stimulus. Slides, video disc frames or real 3-D objects were presented either to the side of the LED or projected to cover the LED at each trial. Presentations to the side of the LED were within 2 degrees of the LED. Presentation was for 1 second after a 0.5 s warning tone. The stimuli were either real 3-D static presentations of an experimenter (or control object), or 2-D slides, or still frames on a video disc. Each stimulus was presented 5 or more times in computer controlled pseudo-random order. In addition, a `no stimulus' condition was also used, where only the

LED and wall could be seen. The inter-trial interval was varied between 0.5 and 5 seconds.

Recording techniques

For each recording session, topical anaesthetic (lignocaine hydrochloride, Xylocaine 40 mg/ml) was applied to the dura and a David Kopf micro-positioner fixed to the recording well. A trans-dural guide tube was inserted 3-5 mm through the dura and a tungsten in glass microelectrode (Merrill and Ainsworth 1972) advanced with a hydraulic micro-drive to the temporal cortex. Position of the cannula was varied using a micro-positioner (David Kopf Instruments) adapted to permit up to 20 degrees of tilt. The target area for recording was area STPa in the anterior part of the upper bank of the STS (which includes areas TPO, PGa of Seltzer and Pandya 1978). The electrode was advanced using a micro-drive (David Kopf 607W) and the depth of each cell recorded noted.

The electrical signals were amplified (Neurolog NL104) and then filtered with a 50 Hz notch filter together with low (300 Hz) and high (20 KHz) pass filters (Neurolog NL125). Spikes from individual cells were discriminated using a threshold voltage window (Modified Digitimer DM130), set for each cell tested and by the visual appearance of the spike on a fast time base oscilloscope. The spikes were converted to TTL signals and these were used to form peri-stimulus time histograms (PSTHs) with 250 bins usually of size 5.2 ms but 5.0 or 4.8 ms in some recordings. The PSTHs had a 1 second post-stimulus period and either 200, 250 or 300 ms pre-stimulus sample periods, depending on the bin size. Data were stored using CED 1401 (Cambridge Electronic Design) and custom software.

Measurement of cell responses

The time at which the shutter became transparent or was fully open was recorded. Subsequent analysis was linked to this, the true stimulus onset time. Neuronal firing rates were measured using standard techniques for a period of 250

ms beginning 100 ms after stimulus presentation. This analysis period was selected because most cells in the STS have latencies of 100-150 ms and because few eye movements occur in this period (see below). A 500 ms sample period was occasionally used for cells with small or late responses. These data were analysed on-line by a microcomputer (AT compatible PC (Hyundai, Dell)).

Eye movement recording

Horizontal and vertical eye movements were monitored using an infra-red corneal reflection system (ACS, modified to allow recording of both signals from one eye) to determine whether any response differences reflected differential patterns of fixation. Differentiating the eye position information allowed assessment of whether speed or velocity of eye movements affected response magnitudes. The output was sampled at the same rate as single cell signals and stored with each trial on the CED 1401 with 8 bit accuracy.

It has been reported that STP cells show differential responses depending on eye movements (Colby and Miller 1986; see Colby 1991 for an example of such a cell). As the effects of eye movements on cell responses were not expressly tested, and eye position was not monitored for all the cells tested, the possibility that some of the responses observed were influenced by eye movements cannot be excluded. However the effects of eye movements are likely to be small, as only 20% (18/90) of cells in STP were found to be related to eye movements. Of these 18 cells, 9 were visual (responding to the onset of the target stimulus) and 4 were visuomotor (firing from target stimulus onset until the saccade was made), while only 5 cells were related exclusively to the saccade (C.L. Colby, personal communication). Thus, only 6% of cells in the present study would be expected to be related solely to saccadic movements and not the stimulus. Further, it is not clear if the cells reported by Colby and colleagues were recorded from the posterior (STPp) or anterior portion (STPa) of STP. Given the large proportion of eye movement related responses in MST and that the input to STPa from MST is

via STPp, one would expect there to be more eye movement related responses in STPp than STPa. For several other reasons given below it is unlikely that the cell responses reported in this thesis reflect differential eye movements.

The experiments described here used a task in which the subject was not required to maintain steady fixation throughout the whole trial period. The monkey performed the LED colour discrimination task with a high level of accuracy, and more importantly obtained multiple rewards by repeatedly licking. The short period during which reward was available meant that the monkey had to be attending the LED from the trial onset in order to lick more than once. Examination of eye position records showed that this was indeed the case, and that fixation was only broken some 400 ms after stimulus onset (see Figures 3.2, 5.2, 6.6, and 7.4). The analysis of the response magnitudes was based on spike counts between 100 and 350 ms (post-stimulus) and therefore during the period of maintained fixation. More than 90% of the cells responding to pursuit eye movements in MT and MST had the eye movement related response starting *after* the onset of the pursuit eye movements (Newsome et al. 1988). This would suggest eye motion 'contamination' of response less likely in the analysis period.

For the few trials where fixation was broken before 350 ms post-stimulus no clear change in either the response latency or response magnitude was observed (see Figures 3.2, 5.2, 6.6, 7.4). Furthermore, as can also be seen in these Figures, even when eye movements were comparable between two stimuli, only one (the preferred) stimulus would give a clear response.

Studies of MST neurons that responded with directional selectivity during pursuit eye movements (where there was no stimulus motion on the retina) also show clear directional responses to retinal motion when the animal was fixating throughout the trial period (stimulus motion but no eye motion). More important the preferred directions obtained under pursuit and fixation tasks were coincident (Erickson and Thier 1991; Komatsu and Wurtz 1988a; Thier and Erickson 1992). Even if the directional selectivity observed in STPa was reflected cell tuning for

direction of pursuit eye movements the findings in MST suggest that the directional selectivity would be the same under stimulus motion with maintained fixation.

It is also relevant to note that many cells recorded in MST do not respond to self-induced retinal motion produced by eye movements (Erickson and Thier 1991). Cells in STP receive a direct input from MST, so they may reflect other response characteristics of MST cells. Indeed it has been argued that there is a trend for cell responses to become less sensitive to self-induced retinal motion at higher in the motion processing hierarchy (Erickson and Thier 1991). It has been shown that STPa cells responsive to motion (and similar to those reported here) are not responsive to equivalent self-induced motion (Hietanen and Perrett 1992). If this is the case then the effects of eye movements on measurements of preferred directions in STP is likely to be small.

It is also relevant that the receptive field size of cells in STPa is very large and typically covering the fovea (Bruce et al. 1981). Similar selectivity for static stimuli at different positions within the large receptive fields has been reported for cells in infero-temporal cortex and STPa (Desimone et al 1984; Gross 1992, Tovee and Rolls 1993). Although the receptive fields for all cells were not expressly checked, of those cells which receptive fields were mapped, similar positional invariance was observed within the STPa. Selectivity for static and moving stimuli was maintained to eccentricities of 10-20 degrees either side of the fovea (Perrett et al. 1989b; unpublished studies Perrett, Harries and Oram). Thus with large receptive fields and positional invariance, difference in eye positions (+/-20 degrees) is unlikely to have effected response selectivity in this study.

In summary these arguments indicate that it is unlikely that the observed selectivity for biological motion stimuli was due to eye movements. First, the available evidence suggests that cell responses in STPa are generally unrelated to eye movements. Second, given the size of STPa cell receptive fields and

positional invariance, any small variation in eye position would not account for differential responses. Third, direct measurements of eye position indicated that differences in eye position/velocity across stimulus conditions were indeed small. Finally and more importantly there was no consistent relation between eye position/velocity and neural responses reported here.

Localization of recording session.

Frontal and lateral X-radiographs were taken of the position of microelectrodes at the end of each recording session.

HISTOLOGICAL AND ANATOMICAL METHODS

After all recording sessions were finished, and any anatomical tracers had been administered, the monkey was sacrificed for histological reconstruction of recorded cell position and anatomical studies. The methods by which this was done can be broken down into 3 stages. These are (i) perfusion fixing and removal of the brain, (ii) sectioning and staining of the brain tissue and (iii) the reconstruction process itself.

Perfusion, Fixing and Removal of the Brain

The monkey was given an anti-coagulant (Heparin, 5000 units for every 1.5 Kg of body weight) then left for approximately 30 minutes to allow the anti-coagulant to pass throughout the body. A ketamine injection (1.0 ml Vetalar) was followed by a lethal intravenous injection of barbiturate (Sagatal). The animal was assessed to be in a deep coma by the absence of a gabella reflex (gently tapping the eye produced no reflexive closing of the eyelid).

Perfusion was performed trans-cardially, using first a pre-fixative wash to flush the blood from the brain, then passing the fixative through the tissue. After

this, the brain was removed and kept refrigerated in phosphate buffer until sectioned and stained. The fixative used was phosphate buffered 4% paraformaldehyde, 0.5% glutaldehyde.

In more detail, the animal's thorax was opened to reveal the heart and a large bore cannula (internal diameter 2 mm) was inserted into the left ventricle . The right atrium was also punctured to allow the exit of the blood / pre-fixative.

After cannula placement the pre-fixative wash (Hartmann's solution, pre-warmed to 37 degrees centigrade) was passed through the animal (approximately 5 litres of pre-fixative for 5-8 Kg animals) with a centrifugal pump to assist the flow. After the pre-fixative wash, the perfusing fluid was changed to the fixative. To ensure good fixation of the brain, the perfusing cannula was pushed gently up through the left ventricle and into the ascending branch of the aorta. Similar quantities of fixative and pre-fixative were used. The brain was then removed from the skull using bone cutters and sunk in successively higher concentrations (10, 20 and 30%) of sucrose solution or 2% Dimethylsulphoxide (DMSO) and 20% glycerol (Rosene et al. 1986). These solutions were designed to avoid freezing artifacts in histology. The dural tissue was left intact and surrounding the brain as much as possible. The isolated brain was then placed in 0.1 M phosphate buffer and refrigerated prior to the histological work.

Sectioning and Staining

The posterior part of the fixed brain tissue was cut away to leave a stereotaxically vertical (coronal) surface. Immediately prior to sectioning, the fixed brain tissue was immersed in pre-cooled isopentane. The temperature of the isopentane was maintained between -70 and -90 degrees centigrade using dry ice (CO_2). The tissue was then mounted in a freezing microtome (Bright Instruments Company Ltd) pre-cooled to -30 degrees centigrade. Coronal sections of 25 microns were taken every 250 microns. During sectioning, photographs were taken of the tissue just before a section was taken. With each photograph the

section number and anterior/posterior position was included, as was a small marker of known length (to allow scaling).

Reconstruction

Reconstruction of electrode position was achieved by reference to the positions of micro-lesions (10 microamp DC for 30 s) made at the end of some electrode tracks which were subsequently identified using standard histological techniques. In 3 monkeys additional markers used in calibration of electrode position were provided by micro-injection of anatomical tracers (horseradish peroxidase and fluorescent dyes true blue and diamadino yellow) at the site of cell recording on 3 recording tracks. For these markers the position of injection, recorded in X-radiographs, could be compared to the anatomical location of injection revealed through normal or fluorescence microscopy. Further probes were inserted into the perfused brain prior to removal from the skull. The position of these probes was also recorded with X-radiographs. Two horizontal alignment probes were also inserted per hemisphere along the full anterior-posterior extent of the brain to allow alignment of each section.

From each frontal and lateral X-radiograph, the electrode tip was measured relative to the interaural plane and midline. The position of the electrode at a known height (25 mm above the interaural plane) was also measured. The 3-D trajectory of each track was calculated from these X-radiograph coordinates. Cell positions along each track were then mapped onto the sections (See Harries & Perrett 1991 for full details). Recently software was developed that allowed entry using a mouse of the photographed brain tissue sections. These maps were stored on PC and allow the supersition of the electrode tracks and cell position.

Breakdown of total number of cells

Table 2.1 gives a breakdown of the number of cells recorded from each monkey and the number of cells included in each chapter.

Table 2.1 Number of cells recorded

MONKEY	Chapter Number					
	3	5	6	7+8	9	Total cells recorded
B	0	36	29	64	0	1300
D	22	48	14	30	4	3182
H	9	4	5	9	5	364
J	13	33	34	58	35	1613
TOTAL	44	216	82	161	44	6549

Table 2.1. In addition to the cells shown, 95 cells from money F were included in the analysis of chapter 5 (directional tuning).

The nature of neurophysiological recordings of the type described in this thesis is such that a collaborative effort is required for data collection. My collaborators for the work presented here have been D. Perrett, and over the years, M. Harries, J. Hietanen, P. Benson, S. Thomas, R. Bevan, H. Ortega, and W. Dittrich. I participated in the collection of the data (i.e. in the laboratory) approximately half the time with the exception of monkeys F, B and the first third of the data from monkey D.

About two thirds of the way through the data collection for monkey D (and for the subsequent monkeys), I wrote the software that was used for data collection, analysis and display. The computerization of data collection formalized and led to the final design of the testing. Without the data collection program the detailed examination of the time course of the neural responses

would not have been possible. I also wrote the software and entered the brain slice data that is used for the histological reconstruction for monkey J.

I have assisted Dave Perrett with 2 operations, and been responsible for maintaining anaesthesia in another one (and an observer in another). I have performed perfusions and helped with the sectioning of the brain for three monkeys.

CHAPTER 3

THE EFFICIENCY OF THE VENTRAL PATHWAY TIME COURSE OF NEURAL RESPONSES DISCRIMINATING DIFFERENT VIEWS OF THE FACE AND HEAD

(Oram & Perrett, 1992, *J. Neurophysiol.*, 68:70-84)

INTRODUCTION

The response of neurons in the inferior temporal cortex (IT) has been associated with the presentation of complex visual stimuli (e.g. Gross et al. 1972; Schwartz et al. 1983; Desimone et al. 1984; Tanaka et al. 1991). In this region and in the anterior superior temporal polysensory area (STPa) of the superior temporal sulcus (STS) cell populations appear selectively tuned to visual stimuli having biological importance (e.g. faces and hands, Bruce et al. 1981; Baylis et al. 1985; Rolls 1984; Perrett et al. 1982, 1989, 1991; Kendrick and Baldwin 1987; Yamane et al. 1988; Hasselmo et al. 1989a,b). The majority of neurophysiological studies within the temporal cortex have focused on the visual basis of selectivity between different visual patterns. The effects of attention (Richmond and Sato 1987; Richmond et al, 1983; Moran and Desimone 1985; Sato 1988, 1989; for review see Desimone et al. 1990) and short term memory (Fuster 1990; Coburn et al. 1990; Riches et al. 1991; Chelazzi et al. 1993) on visual responses have also been studied in this area. Little is known about the dynamics of neural responses, or the dynamics of the discrimination between responses.

Studies of neural response dynamics have so far been restricted to the way the temporal modulation of firing rate over several hundred milliseconds relates to the stimulus pattern (Richmond et al. 1987; Richmond and Optican 1987; Optican

and Richmond 1987; Gawne et al. 1991a,b; McClurkin et al. 1991a,b) or the possible functional role of oscillations (Gray et al. 1989; Engel et al. 1992a,b; Singer 1990a,b). Different stimulus patterns not only affect the mean discharge rate sampled over an *extended* period of time but stimulus pattern differences can also be related to different temporal patterns of cell activity (e.g. maintained activity, a single transient response or two or more oscillating peaks of activity; Optican and Richmond 1987; McClurkin et al. 1991b; Gawne et al. 1991a). Recently the results relating the temporal pattern of cell responses to pattern differences has been questioned (Tovee et al. 1993), as has the possible role of oscillations (Young et al. 1992; Tovee and Rolls 1992; Ghose and Freeman 1992; Singer 1993).

Monkeys can be trained to discriminate behaviourally whether or not a complex visual pattern has been seen before (i.e. whether it is familiar or novel) within 450 ms (Perrett 1981, unpublished thesis). Since cells in the anterior thalamus discriminate between novel and familiar visual patterns within 150-200 ms during performance of this recognition task (Rolls et al. 1982; Perrett 1981, unpublished thesis) at least half the reaction time is taken up with the generation of motor commands and contraction of muscles. This indicates complex pattern recognition can occur within a short time, even given the sluggish performance of the neural computing elements. In such a short time, only a limited number (20-30) of processing steps or stages can occur between retinal output and visual recognition of stimuli as familiar. Indeed Thorpe and Imbert (1989) have argued that perhaps as few as 10 synapses occur between the retinal photo-receptors and cells in the STPa which are selectively responsive to the sight of faces.

Recordings of cells from the temporal cortex offer an opportunity to examine the time course both of responses per se and of the discrimination between responses to different complex visual input patterns. These parameters are of interest since they can be compared to the performance predicted by models

of visual processing, particularly those with networks of neuron-like elements, for example parallel distributed processing (PDP) models (see following chapter).

This chapter focuses on whether or not the firing rate of cells, during the initial part of the response, exhibits discrimination between visual patterns. This issue relates to predictions from PDP models (see following chapter) and is orthogonal to whether or not the temporal pattern of discharge over an extended time period contains information about the stimulus nature, though, if there were no information available to discriminate between stimuli during the initial response of cells, then models of processing need to take account of the response pattern over extended of periods of time.

METHODS

The single cell recordings were made using standard chronic implant techniques (see Perrett et al. 1985, 1991) from 3 awake behaving rhesus monkeys [2 male (D,H) weight 7.6 Kg each and 1 female (J) weight 4.8 Kg]. The standard behavioural and recording techniques were used.

Stimuli

Up to 8 views of the head in the horizontal plane and a range of control objects (matched to the head for approximate size and varying in colour and texture) were used. Stimulus presentation was for 1 second after a 0.5 s warning tone. The stimuli (heads and control objects) were either real 3-D static presentations of an experimenter (or control object), or 2-D slides, or still frames on a video disc. Each stimulus was presented 5 or more times in computer controlled pseudo-random order. In addition, a `no stimulus' condition was also used, where only the LED and wall could be seen. The inter-trial interval was varied between 0.5 and 5 seconds.

Data analysis

Classification of cells was based on the mean cell response (spikes per second) over a period of 250 ms starting 100 ms after stimulus onset. The dynamics of neural responses (the subject of the present study) were analysed only for cells that were classified as viewer-centred. Briefly, viewer-centred cells were defined as those where ANOVA and post-hoc testing (Protected Least Significant Difference, PLSD: Snedecor and Cochran 1980) of the responses showed at least one view of the head was greater ($P < 0.025$, 1-tailed test) than both spontaneous activity (S/A) and control levels but another head view gave a response that was statistically indistinguishable from S/A and/or controls (see Perrett et al. 1991 and Appendix).

For each cell, responses to head views were categorized into three groups: Best (response to the most effective stimuli), the Worst (response to the least effective stimuli) and the Mid (producing a response mid way between these). The Best, Mid and Worst categories were each defined as a 20% sub-range of the full range of responses to different head views tested: Worst = 0-20%, Mid = 40-60% and Best = 80-100%. This gave Best and Worst responses for all cells. For some cells no head view tested gave a mean response which fell into the Mid range. Within each category the responses were, therefore, all of a comparable strength relative to the cells' maximum and minimum responses. If no evidence were found for relationships between the particular stimuli (e.g. view of the head) and response characteristics, then each category would contain a set of responses that was only determined by their relative strength. This process therefore allows a greater number of individual cell responses to be directly compared. Given the sparsity of cells selective for different views of the head (only 5% of visual cells in the STS, Perrett et al. 1991: < 1% of all recorded cells), this grouping was necessary to allow the analysis of the time course of pattern discrimination in a population of cells. This grouping therefore allowed the response levels of

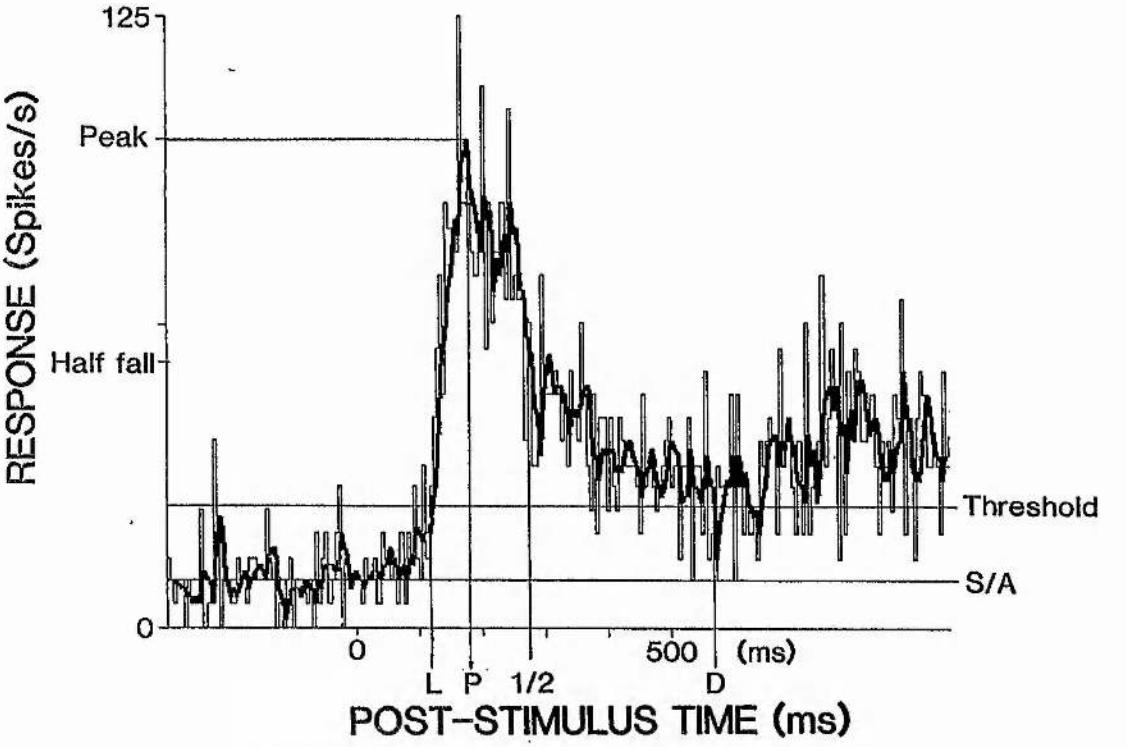


FIGURE 3.1. MEASUREMENT OF NEURAL RESPONSE PARAMETERS FROM PERI-STIMULUS TIME HISTOGRAM. Magnitude of response from several trials (typically n = 5) is displayed relative to stimulus onset (0 ms). Thin histogram bars give response rate above spontaneous activity (S/A) in successive 5 ms time bins. Thick line show response magnitude smoothed using a running 3 time bin average. Threshold of response = the upper limit of the 95% confidence interval of S/A. Latency of response onset (L) = first of 3 consecutive 5 ms bins, each exceeding the threshold. End of response (D) = time when smoothed response was less than threshold. The Peak firing rate = maximum of the smoothed response rate (at time = P). Half-fall (1/2) = time for smoothed response to decay from peak to half peak response. `Rise time (Time to peak) = P - L; Decay time = D - P, Response Duration = D - L.

different relative magnitude to be compared across cells independently of the particular pattern preferred by the individual cells.

The responses on different trials (typically 5 or 10) were averaged within each response category for each cell (on a bin by bin basis) after aligning the recorded stimulus onset times. This created a PSTH for each response category for each cell. A number of response measurements were made from these PSTHs (see Figure 3.1). In addition to those shown, the mean firing rate of the cell was calculated during (a) the pre-stimulus period, (b) the first, second and fifth 100 ms of the response after the estimated cell response latency and (c) the period 150 to 50 ms before the end of the sampling period. This last measure, although not linked to the response onset gave an estimate of the final or stable firing rate. The estimate of the pre-stimulus activity in each category was taken from the averaged trial responses and regarded as the spontaneous activity (S/A). The latency of each cell was taken as the first of 3 consecutive time bins where the mean response was in excess of the 95% confidence interval of S/A. If more than one category (Best, Mid or Worst) for a particular cell gave a different latency estimate, the shortest was taken as the cell response latency.

A sliding average of a 3 bin window (nominally 15 ms) was used to estimate all other measures of activity and response time course (see Figure 3.1). The centre bin of the sliding window was used to define the time at which spike frequency values reached given criteria.

The values of the firing rate measurements were expressed as the rate above S/A and then a further 9 parameters were calculated. For the four 100 ms sample periods, the firing rates were expressed as a percentage of the peak response. The discrimination between the Best and Worst categories was also calculated for the peak response as well as over the four 100 ms samples. The discrimination was expressed as $100 * (R_B - R_W) / R_B$, where R_B is the mean response level above S/A in the Best response category and R_W is the mean

response level above S/A in the Worst response category. If the Best response was within 1 spike per second of S/A the discrimination measure was not calculated.

Population response. PSTH profiles in the Best, Mid and Worst response categories for the entire population of sampled cells were obtained first (1) by normalizing the response magnitude of each cell to the magnitude of the difference between the peak response (measured over 5 ms) of the Best category and S/A, then (2) averaging response rate in each time bin across all cells and finally (3) renormalizing as step (1). This procedure meant that each cell contributed equally to the population estimate.

Average Cell response. For each cell the Best, Mid and Worst responses were shifted by the difference between the cell's estimated response latency and 100 ms. This allowed the responses of different cells to be synchronized to provide an estimate of the time course of the 3 levels of response of an Average Cell.

To determine the statistical efficiency of discrimination between different stimulus categories, the responses in each time bin to each of the three response categories were taken from all contributing cells and subjected to a 2-way ANOVA, with response category being a fixed factor and cell a random factor. For this analysis, only data collected with the same bin size (5.2 ms) were used. Firing rates above S/A of each response category for each cell were used without normalizing the magnitude of response. This reduces the potential problem of having unequal variance between the three categories. For example, if all the individual cells reached peak response magnitude of the Best category in the same time bin, then normalization would lead to all values in the Best category for that bin having value unity with zero variance. Discrimination analysis was also performed using transformed data to correct for the standard deviation being proportional to the mean (logarithmic transform), and for the variance being proportional to the mean (square root transform). The F ratio for each time bin

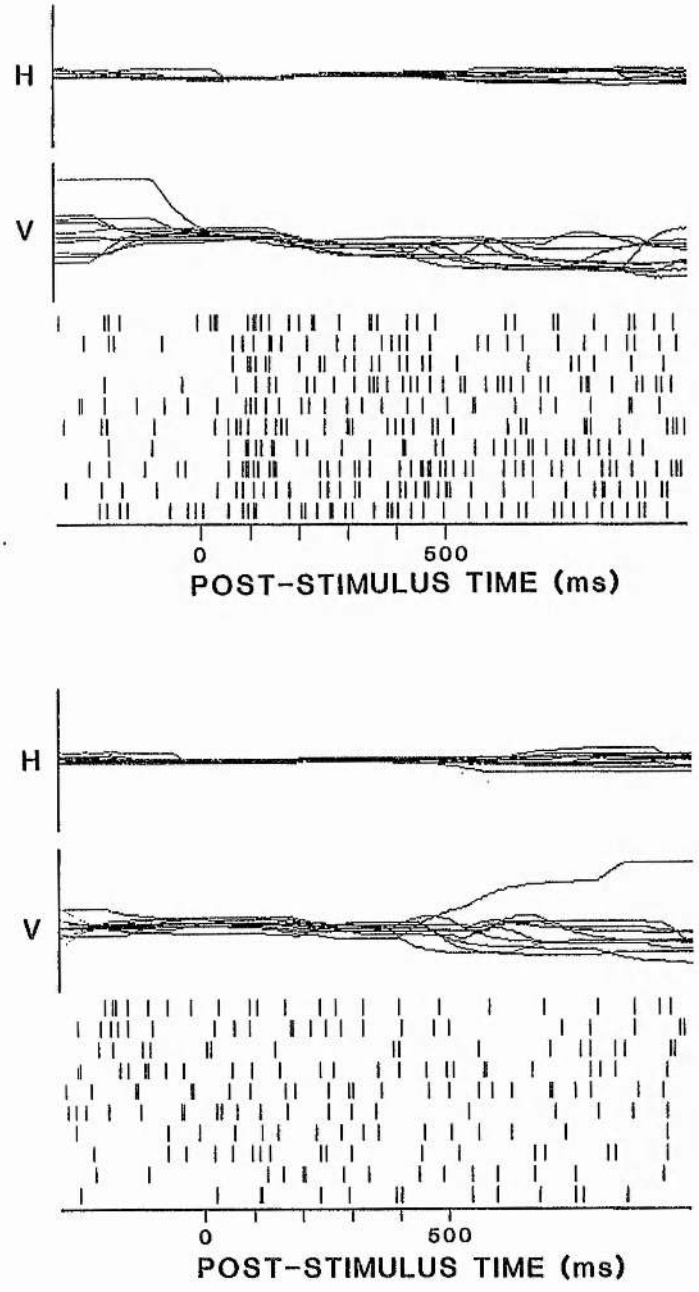


FIGURE 3.2. EFFECT OF EYE MOVEMENTS ON RESPONSES. **UPPER: EYE POSITION AND RESPONSES DURING EFFECTIVE STIMULI.** Eye movement traces and cell responses during 10 presentations of an effective stimulus (Best category responses). **LOWER: EYE POSITION AND RESPONSES DURING INEFFECTIVE STIMULI.** Eye movement traces and cell responses during 10 presentations of a non effective stimulus (Worst category responses). Full scale deflection (FSD) of horizontal eye movements (H) = 40 degrees, FSD of vertical trace (V) = 30 degrees. For both categories of response, the neural activity associated with each pair of eye traces is shown in rastergram form.

was taken as a measure of discrimination between stimuli of different degrees of effectiveness that the population or Average Cell showed over that time period.

RESULTS

110 cells were classified as having selective viewer-centred responses to the sight of the head. This was approximately 5% of the total number of cells recorded in the STS upper bank (Perrett et al. 1991). Data were available in a suitable form for analysis of response time course for 44 of these cells (22 from monkey D, 9 from H and 13 from J).

To investigate the possibility of differences in the response measures between monkeys, 1-way ANOVAs were performed on all parameters listed in Table 3.1. The results indicated no significant difference between monkeys for all parameters ($P > 0.17$, all comparisons). Parameter measurements for individual cells were therefore pooled across monkeys.

All cells contributed to the Best and Worst estimates of the population response but 16 cells were not tested with a view of the head that gave a response in the Mid range. The comparisons made between the Best, Mid and Worst categories were therefore performed using data from 28 cells, of which 22 had a bin size of 5.2 ms.

Eye movements

Fixation patterns were found to be similar independent of the stimulus used. The subjects showed good fixation of the LED position, starting before stimulus onset and lasting some 400-500 ms post-stimulus onset (Figure 3.2). Neuronal responses to an effective stimulus were comparable across repeated presentations (Figure 3.2a) but were absent to presentations of a non-effective

Table 3.1

(a) Summary of Best Category Response Parameters

Parameter	MEAN	RANGE		N
Timing				
Latency (ms)	119.05	69.2 to	212.8	44
Rise Time (ms)	58.16	4.8 to	239.2[A]	44
1/2 Fall Time (ms)	40.00	20.8 to	93.6	44
Decay Time (ms)	93.44	20.8 to	525.2	43[B]
Duration (ms)	112.46	14.4 to	561.6	43[B]
Firing Rates				
S/A (Spikes/Sec)	8.62	0.3 to	23.6	44
Peak (Spikes/Sec)	115.13	51.3 to	256.4	44
1st 100ms (Spikes/Sec)	66.88	24.3 to	135.4	44
2nd 100ms (Spikes/Sec)	48.10	2.0 to	114.6	44
5th 100ms (Spikes/Sec)	28.46	7.9 to	91.1	44
End 100ms (Spikes/Sec)	24.73	6.0 to	98.2	44
Normalized Response Magnitudes[C]				
1st 100ms : Peak (%)	54.30	28.5 to	72.8	44
2nd 100ms : Peak (%)	38.00	0.0 to	72.3	44
5th 100ms : Peak (%)	18.75	1.2 to	51.4	44
End 100ms : Peak (%)	15.08	-2.4 to	51.5	44

(b) Summary of Discrimination Measures [(Best - Worst) / (Best-SA)]

Peak Disc (%)	61.03	3.0 to	106.0	44
1st 100ms Disc (%)	72.55	7.8 to	116.2	44
2nd 100ms Disc (%)	86.65	21.9 to	166.0	43[D]
5th 100ms Disc (%)	83.77	-170.0 to	200.0	44[E]
End 100ms Disc (%)	66.01	-280.0 to	268.4	42[D,E]

[A]The longest rise time seen (239.2 ms) was exceptional, the next two longest being 182.0 ms and 140.4 ms. [B]For one cell decay time and duration measures were greater than the sample period. [C]Normalized response magnitudes were calculated as (Firing rate - S/A) / (Peak - S/A). [D]Discrimination measures were not calculated if the Best response level was within 0.1 spikes/second of S/A. [E]Negative discrimination measures resulted when the firing rate in the

Worst category was greater than the Best, or in 1 case (End 100ms discrimination), when the Best firing rate was less than S/A but greater than the Worst firing rate.

stimulus (Figure 3.2b). The similarity of the eye position during trials to both stimulus types shows that the responses were stimulus driven and not due to differences in fixation patterns.

Relationship between response and view

The visual properties of the head vary between different views. For instance the back view can be thought of as an almost uniform blob (with low spatial frequency) with details of the hair texture (with high spatial frequency), whereas the face view has more visual components (e.g. eyes, nose, mouth, hairline etc) and has a greater range of spatial frequencies. Low and high spatial frequency information is carried by different populations of cells in early visual processing. A check was therefore made as to whether any correlation was present between the cell's most effective view of the head and any of the response parameters measured. The most effective view was determined by regression of a second order cardioid equation on the responses to all views tested (see Perrett et al. 1991 for details). The resulting angle estimates were expressed as degrees of rotation away from full face view, independent of direction of rotation (i.e. left and right profile = 90 degrees).

No correlation was found between the most effective view and any of the response parameters. This lack of correlation was found to be the case for all three response categories. A similar lack of reliable correlation was found when the angle of rotation was measured from profile views. The absence of systematic variation of response time course, response magnitude and discrimination measures with view indicates that the dynamics of response are similar for cells selective for different head views.

Time-course of response to effective stimuli

The measured response parameters for the Best category of response are listed in Table 3.1. The response of the cell population showed a very rapid and

Table 3.2

Comparison of the Best, Mid and Worst categories of response

Parameter	Best	Mid	Worst	F	df	p
Timing (ms)						
Latency	111.2	120.4	127.7	1.195	2/22	0.322
Rise Time	52.8	64.4	32.6	2.827	2/48	0.069
1/2 Fall Time	40.6	30.3	24.1	7.512	2/46	0.002
Decay Time	83.6	36.5	26.1	18.424	2/44	0.000
Duration	128.6	63.3	50.6	5.501	2/20	0.012
Firing Rates (Spikes/S)						
S/A	9.4	8.9	9.3	0.203	2/54	0.817
	(2.9	2.7	2.9	0.989	2/54	0.378)
Peak	115.8	82.4	51.2	84.793	2/54	0.000
	(10.6	8.8	6.7	85.730	2/54	0.000)
1st 100ms	66.1	43.6	24.3	78.281	2/54	0.000
	(7.9	6.4	4.6	107.173	2/54	0.000)
2nd 100ms	45.1	31.8	16.0	51.300	2/54	0.000
	(6.4	5.3	3.7	71.084	2/54	0.000)
5th 100ms	27.7	19.7	13.1	16.977	2/54	0.000
	(5.1	4.2	3.2	19.577	2/54	0.000)
End 100ms	25.6	18.1	12.4	11.913	2/54	0.000
	(4.8	4.1	3.2	17.416	2/54	0.000)
Normalized to Peak Response Magnitudes (%)						
1st 100ms	53.3	47.0	34.1	7.940	2/54	0.001
2nd 100ms	34.7	33.2	13.5	9.619	2/54	0.000
5th 100ms	17.7	14.5	7.7	2.629	2/54	0.081
End 100ms	15.0	13.9	4.8	2.847	2/54	0.067

Means for each parameter under each category are listed, with the resulting variance ratio (F), degrees of freedom (d.f.) and the probability (p) of the values being statistically the same. For firing rates, the analysis was also carried out using square root transform of the data (values in parentheses). S/A, spontaneous activity.

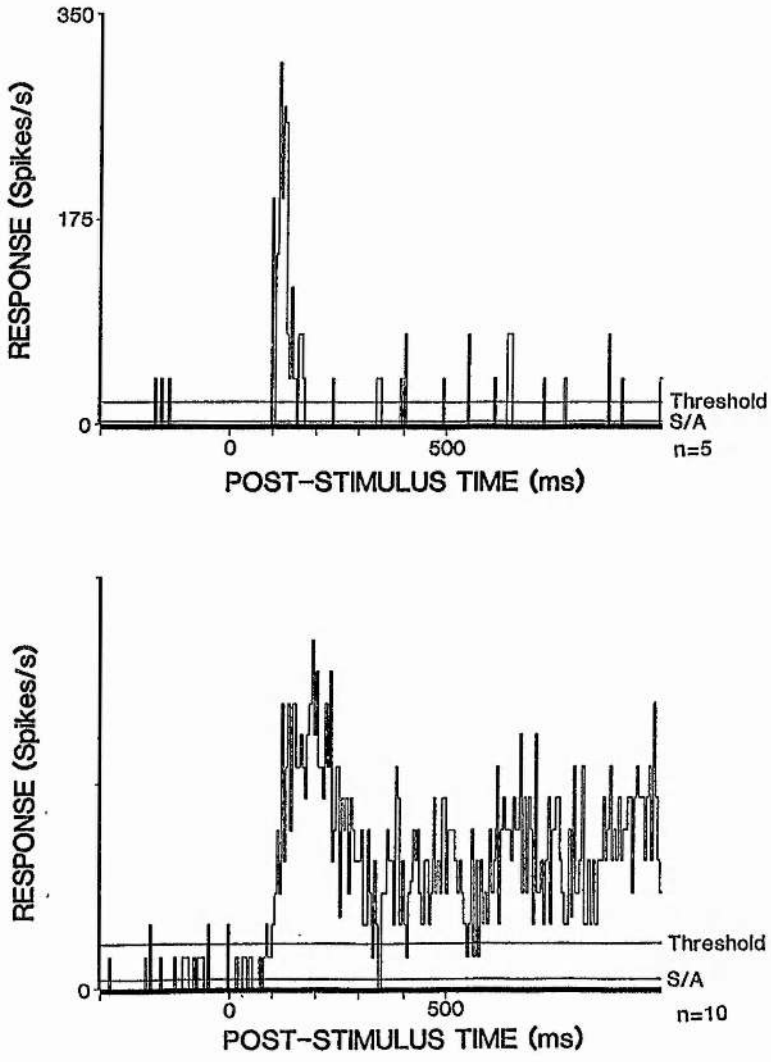


FIGURE 3.3. FASTEST AND SLOWEST RESPONSE DECAYS. **UPPER: RAPID RESPONSE DECAY.** Peri-stimulus time histogram of the Best category of responses of one cell with a very short transient response (number of contributing trials to PSTH, n=5). Horizontal line = threshold (see Figure 3.1). **LOWER: SLOW RESPONSE DECAY.** The Best category of response of a cell exhibiting the slowest rate of decay of response. The smoothed response (not shown) did not fall below threshold during the sample period (n=10).

clear increase in firing rate to the preferred views, followed by a slower decline in firing rate. Evidence of the rise and fall in firing rate during the response was seen for all cells. The rapid rise in firing rate was a feature across most cells (rise time ranging from 5 to 239 ms, mean 58 ms). For some cells the decay of firing rate was very rapid and complete, the initial part of the response being a short, intense, transitory burst. Such transient activity is illustrated for one cell in Figure 3.3a. After the initial peak in firing rate, the response rate of this cell remained slightly increased over the S/A for the duration of the stimulus presentation. Most cells showed a slower rate of response decay but still showed evidence of an initial transient burst (e.g. Figure 3.3b). A gradation of decay time course between these two examples was found (decay time ranging from 21 to 525 ms). All cells showed evidence of a response maintained above spontaneous activity up to the end of the sampling period.

Best, Mid and Worst category responses

Comparison of the responses to three sets of stimuli with different effectiveness allows the time course of discrimination in cell responses to different head views to be defined. As already noted, the Best responses show an initial burst and decay to a steady value which is greater than the pre-stimulus activity.

In contrast there was a greater variety in shape of the Worst response. Some cells showed no response at all to these stimuli. The most frequent pattern was a transient response similar in shape to the Best category but greatly reduced in magnitude. Only a few cells showed evidence of inhibition to the Worst view in the early part of response (i.e. within 100 ms of the latency estimate of the Best response, see Figure 3.4a).

Since the Worst category actually contained a range of response magnitudes for each cell (0-20 % of response range to different head views) any

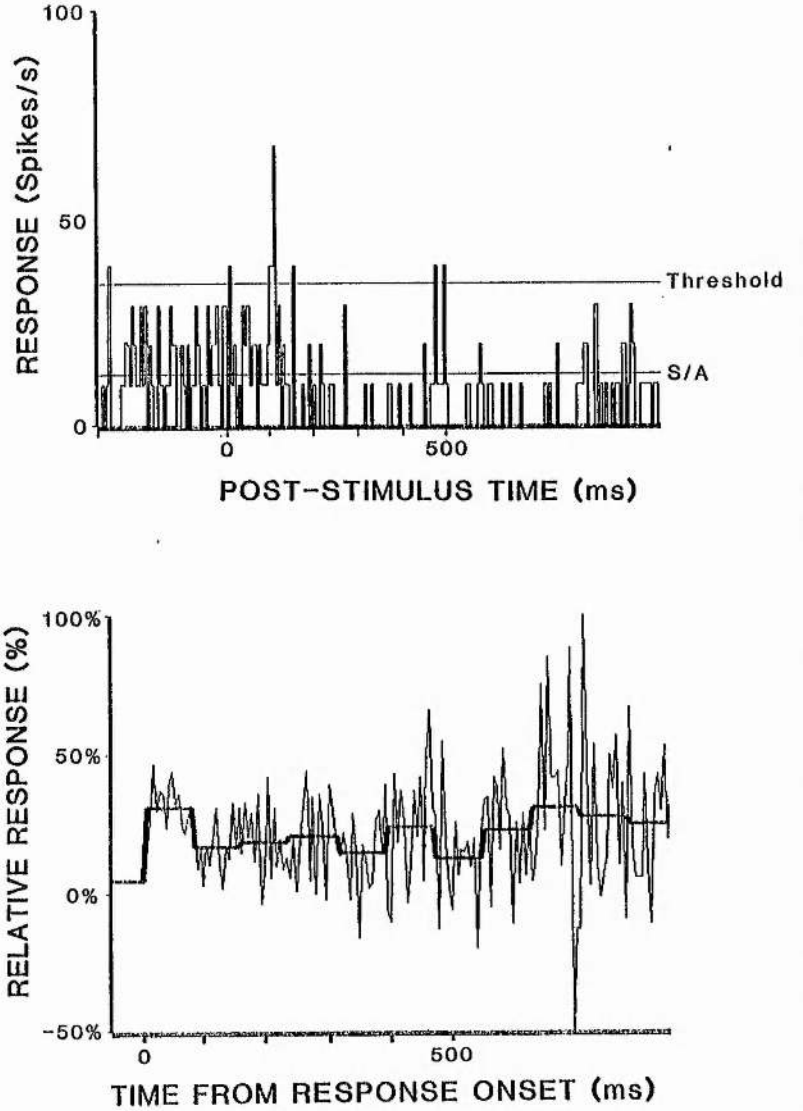


FIGURE 3.4. INHIBITION TO INEFFECTIVE STIMULI. **UPPER: INHIBITORY RESPONSE ON A SINGLE CELL.** One of the few cells to show inhibition below S/A in the Worst response category. **LOWER: RELATIVE MAGNITUDES OF THE WORST AND BEST RESPONSES.** The mean of 22 cells' Worst responses is expressed as a percentage of the Best category responses after synchronization of response latencies. Thin line: 5.2 ms bins. Thick line: 80 ms bins. The percentage rapidly increases and remains relatively constant at 31% for the first 80 ms, then shows a significant ($t_{[30]} = 4.39$, $P < 0.0005$) fall for the second 80 ms period (mean = 17%). Thus competitive inhibition is slow to emerge and not apparent for the first few tens of milliseconds.

sign of inhibition might have been masked by averaging together weak excitatory responses with inhibitory responses. The data were therefore re-analysed using only trials with the one (least effective) view of the head that produced the lowest response. Expressing the firing rates relative to S/A, the peak firing rate to the least effective view tested for the 44 cells was substantially above S/A (mean = 51.8, S.E.M. = 4.78, range -1.7 to 125). The mean firing rate was also greater than S/A during the first 100 ms of the response (mean 15.7, S.E.M. = 2.71, range -3.5 to 68.7) and during the second 100 ms of the response (mean = 5.51, S.E.M. = 1.77, range = -12.3 to 40.5). During the fifth 100 ms period of response and the final 100 ms of the sample period, the mean firing rates were statistically indistinguishable from S/A (fifth 100 ms period: mean = 3.9, S.E.M. = 1.87, range = -17.6 to 48.6; final 100 ms period: mean = 3.75, S.E.M. = 1.64, range = -7.8 to 42.5).

Although the population mean firing rate to the least effective view was never significantly below S/A, the negative firing rates indicate that inhibition was potentially present for some cells. Due to the low S/A of these cells (mean = 9 spikes/s), statistical detection of inhibition for any one cell was unlikely. This would be especially true if any inhibition was achieved by mechanisms similar to that described in the cat V1 (Douglas and Martin 1991), where inhibition acts to stop re-excitation from other cortical neurons. The number of cells showing a firing rate numerically below S/A was 1 during the peak estimate, 6 during the first 100 ms of response, 16 during the second, 18 during the fifth and 19 during the final 100 ms period examined. Cells responding at rates numerically less than S/A may not reflect real differences from S/A and may simply be due to sampling error. This was probable for several of the cells enumerated above, since differences were often small (0.1 - 2.0 spikes per second). A more detailed examination of possible inhibition or suppressive effects is given later.

The time course and response amplitude parameters were compared across the different categories of response using a 2-way ANOVA analysis for the 22

cells where head views were tested that fell into all three categories of response. The results of the comparisons between the three categories of responses are shown in Table 3.2. Not surprisingly, the average firing rate was significantly different across the three response categories during the early part of response. It is more notable that the response differences are maintained throughout the sample period (for at least 800 ms after response onset).

The latency and rise time of the responses did not vary across response category. [This suggests that the information about the different types of stimulus (head views) arrives at the same time and has a similar temporal distribution.]

The nature of the decay from the peak firing rate to the steady state firing rate does, however, show a significant difference between the response categories. Post-hoc testing indicates that the 1/2 fall time of the Best response was significantly longer than that for the Mid or Worst categories of response ($P <$ 0.02 Best vs Mid; $P < 0.0005$, Best vs Worst). A different rate of response decay was also indicated in two other measures. The decay time (time from the peak of the response to a 3 time bin mean where the response is statistically equal to S/A) was also different across categories (Best > Mid, Worst, $P < 0.0005$ each comparison), as was the response duration (time from response onset to a 5 ms bin where response equals S/A, $P < 0.005$ each comparison).

Further evidence for different rates of decay during the early part of the response (100-300 ms post-stimulus) is apparent in Table 3.3 from the comparison of normalized response magnitude (expressed as a percentage of the peak response in each category). This latter calculation shows that the rate of decline is not linearly related to the peak response rate. It is progressively faster for lower peak rates. This is consistent with a winner-take-all type of network with competitive inhibition between network nodes.

Time course of relative response magnitudes

As already noted there was little sign of inhibition to sub-optimal stimuli within the cell population. Given that there was some evidence (e.g. Figure 3.4a), a closer examination was performed by comparing the relative magnitude of responses to Best and Worst response categories. This comparison can potentially reveal suppressive effects on cells responding sub-optimally even if the suppression does not result in inhibition relative to S/A. Such suppressive effects could derive from a variety of sources including competitive inhibition from cells responding optimally.

Figure 3.4a shows a peri-stimulus time histogram of the response of one cell to the Worst category stimuli. The cell latency was estimated as 100 ms from the Best category response. Note there is an initial rise in firing rate lasting some 20 to 25 ms, with inhibition becoming evident some 40 ms after the cell's latency to effective stimuli. This cell was exceptional both in showing clear inhibition and in showing inhibition shortly after response onset. This pattern of a small, short duration excitatory peak was typical of responses to the Worst category, even in cells that showed evidence of inhibition. Inhibition, when seen, was typically 60 - 150 ms after the estimated response latency.

The magnitude of the response in the Worst category is expressed relative to the response in the Best category in Figure 3.4b. The only systematic reduction in the relative response apparent at the 5 ms time bin resolution (thin line) occurs 80 ms after response onset. The decrease in relative response magnitude of the Worst category between the first and second 80 ms periods is highly significant ($t_{[29]} = 4.39$, $P < 0.0005$). To emphasize this decrease, the graph also plots relative firing in 80 ms time bins (thick line). This analysis does not indicate the presence of competitive inhibition between cells analyzing different perspective views of the head in the STPa during the early part of their response. Late inhibition was also apparent in the decay of the response, which was significantly faster for weaker response categories (Table 2).

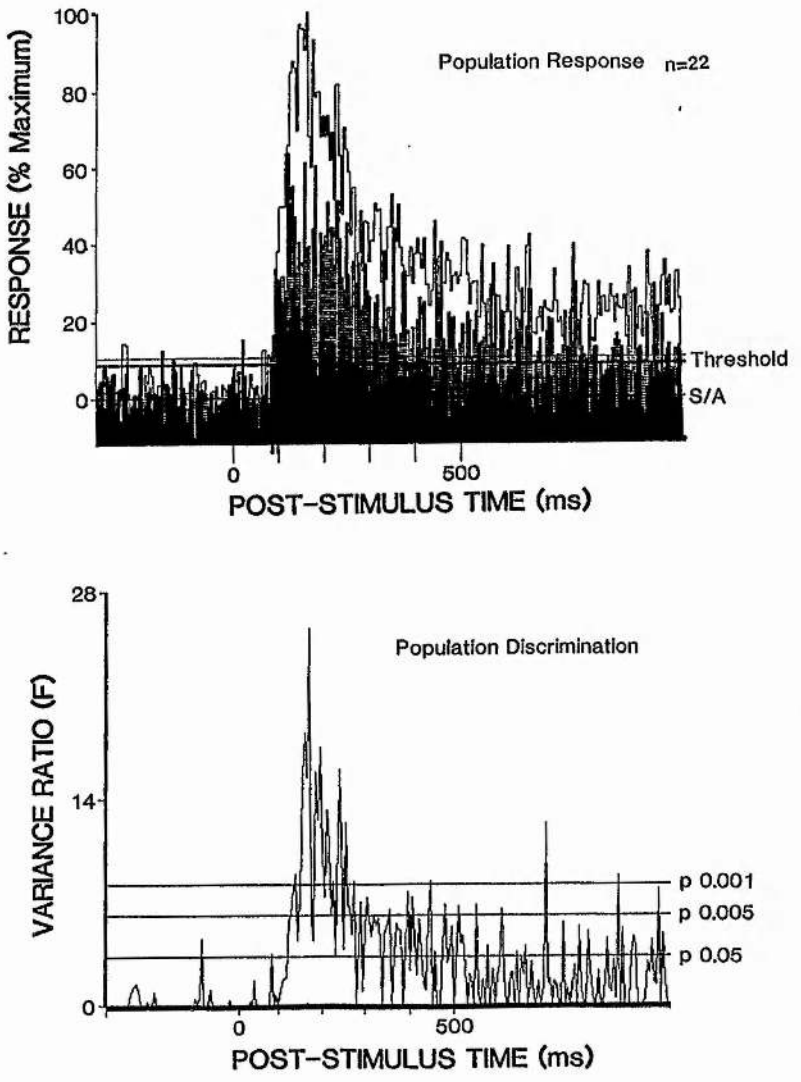


FIGURE 3.5. CELL POPULATION RESPONSE AND DISCRIMINATION BETWEEN HEAD VIEWS. UPPER: COMBINED RESPONSES OF 22 CELLS TO THE THREE CATEGORIES OF STIMULUS EFFECTIVENESS. Clear bars (tops only plotted) = the Best category, hatched bars = the Mid category and solid bars correspond to the Worst category of responses. Firing rate is expressed as a percentage of peak response in Best category. LOWER: STATISTICAL EVALUATION OF RESPONSE DISCRIMINATION. Probability levels of 0.05, 0.005 and 0.001 shown. Discrimination between stimuli first reaches significance ($P < 0.05$) for the cell population at 116 ms after stimulus onset.

Vogels and Orban (1991) showed that the variability of neural responses in primate V1 increased with time lapsed from response onset. Similarly the variability in firing rate of STPa cells assessed with small time bins (thin line) appears to increase with time. This the increase is not apparent in the initial stages of the response but develops slowly over hundreds of milliseconds. Thus any early discrimination seen within 100 ms of response onset is not due to low response variability allowing statistical detection of a small relative difference in firing rates. During the early part of response (first 100 ms), firing rate is high, discrimination is clear (Figure 3.6) and variability is low (Figure 3.4). During the late part of response, firing rate is lower, discrimination is poor and variability is high.

Population response

For the 22 cells with a 5.2 ms time bin found to produce a response in each of the three response categories the amplitude normalized responses (S/A = 0%, Peak from Best = 100%) were averaged by response category to produce the population PSTH response profiles (Figure 3.5a). The parameters as measured for the population responses are shown in Table 3.3a. The increased rise time and 1/2 fall time compared with the mean of the individual cells can be explained in part by the differing response latencies and in part by the reduced variability of firing rate (assessed on a 5 ms basis) when responses are averaged over cells.

A 2-way ANOVA (fixed factor = response category, random factor = the 22 cells) was performed between the three categories of response each of the 250 time bins. The resulting F ratio values, where the rank ordering of the responses was Best > Mid > Worst, are plotted against time in Figure 3.5b. The discrimination between stimuli achieves statistically very reliable and stable levels at 116 ms after stimulus onset. The latency of response of this cell population (n=22) occurs 90 ms after stimulus onset.

Table 3.3

Parameter Values of the Population and Average Cell Responses

	(a) Population			(b) Average Cell		
Parameter	Best	Mid	Worst	Best	Mid	Worst
Timing (ms)						
Latency	95.2	90.0	105.6	100.4	100.4	105.6
Rise Time	62.4	36.4	31.2	57.2	31.2	36.4
1/2 fall Time	119.6	166.4	46.8	130.0	140.4	52.0
Decay Time	842.4	603.2	57.2	842.4	306.8	57.2
Duration	904.8	639.6	88.4	899.6	338.0	93.6
Firing Rate (%)*						
Peak	95.6	56.4	32.4	89.8	53.3	32.7
1st 100ms	69.8	37.6	20.5	72.5	41.8	21.2
2nd 100ms	58.1	36.8	10.9	53.7	33.4	10.6
5th 100ms	23.9	14.4	4.9	24.7	13.4	4.7
Last 100ms	22.8	12.5	4.6	23.5	11.7	5.6
Normalized to Peak Response Magnitudes (%)						
1st 100ms	73.0	66.7	63.3	80.7	78.4	64.8
2nd 100ms	60.8	65.2	33.6	59.8	62.7	32.4
5th 100ms	25.0	25.5	15.1	27.5	25.1	14.4
Last 100ms	23.8	22.1	14.2	26.2	22.0	17.1

Values of response parameters for 3 levels of stimulus effectiveness are listed for (a) the cell population and (b) the Average Cell after synchronisataion of response onsets. *The normalization of firing rates set S/A = 0 and peak for each cell to 100. The values are expressed as percentage of the maximum possible firing rate (100).

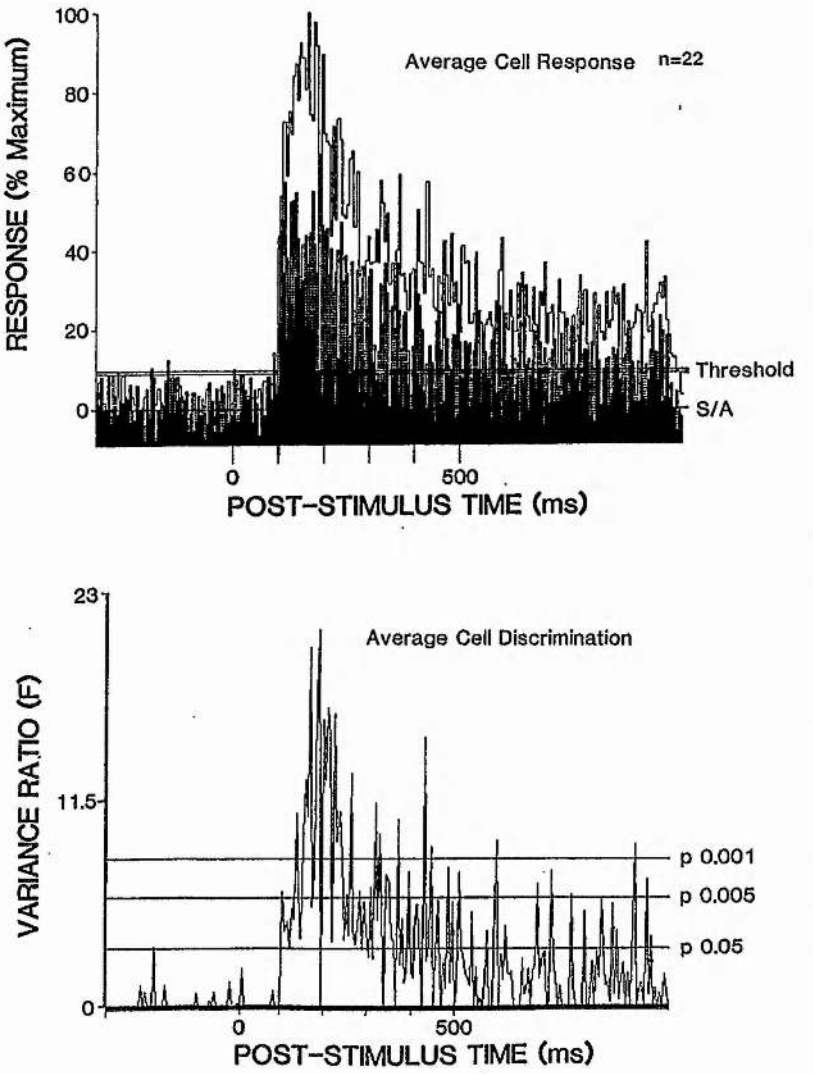


FIGURE 3.6. AVERAGE CELL RESPONSE AND DISCRIMINATION. UPPER: AVERAGE MAGNITUDE OF RESPONSES FOR 22 CELLS TO THREE CATEGORIES OF STIMULUS EFFECTIVENESS. The response latencies of contributing cells were synchronized to 100 ms (post-stimulus). Best, Mid, Worst categories denoted by clear, hatched and solid traces respectively. LOWER: STATISTICAL EVALUATION OF RESPONSE DISCRIMINATION BETWEEN THE 3 RESPONSE CATEGORIES (F ratio computed for each time bin). Discrimination between stimuli reaches significance ($P < 0.05$) within 5 ms of response onset.

A more accurate estimate of the population latency was obtained by examining the responses of all cells selective for head views (n=44). The latency of this larger cell population was 80 ms, with discrimination between Best and Worst occurring at 95 ms.

Average Cell response

Figure 3.6a shows the average response from the 22 cells (each with three response categories measured), after synchronization of response onsets to 100 ms post-stimulus. It should be noted that the method of response onset synchronization was performed per cell and not per response category for each cell. Thus in principle the different categories of response could occur at different latencies in the Average Cell assessment. The Best, Mid and Worst responses, however, occur at the same latency. As can be seen, the shape obtained is qualitatively similar to the population PSTH profiles, except for a faster rise time. The response parameters measured from these PSTH profiles are shown in Table 3.3b.

It is important to note that the latency of onset for each of the three levels of response that has been detected for the data combined from 22 cells does not occur in advance of 100 ms. With a conservative latency assessment a small (possibly indiscriminate) response might have started in advance of the detected response onset. This would become evident in the average.

Discrimination between different levels of response and hence different types of stimuli was analysed by ANOVA of firing rate in different time bins across the 22 cells. This statistical evaluation of the discrimination is illustrated in Figure 3.6b. It is apparent that cells on average discriminate at statistically significant levels within the first 5 ms of response onset.

Variance in firing rate can be correlated with the response strength (Vogels and Orban 1991). A suitable transformation $[(\text{firing rate} + 1)^{1/2}$,

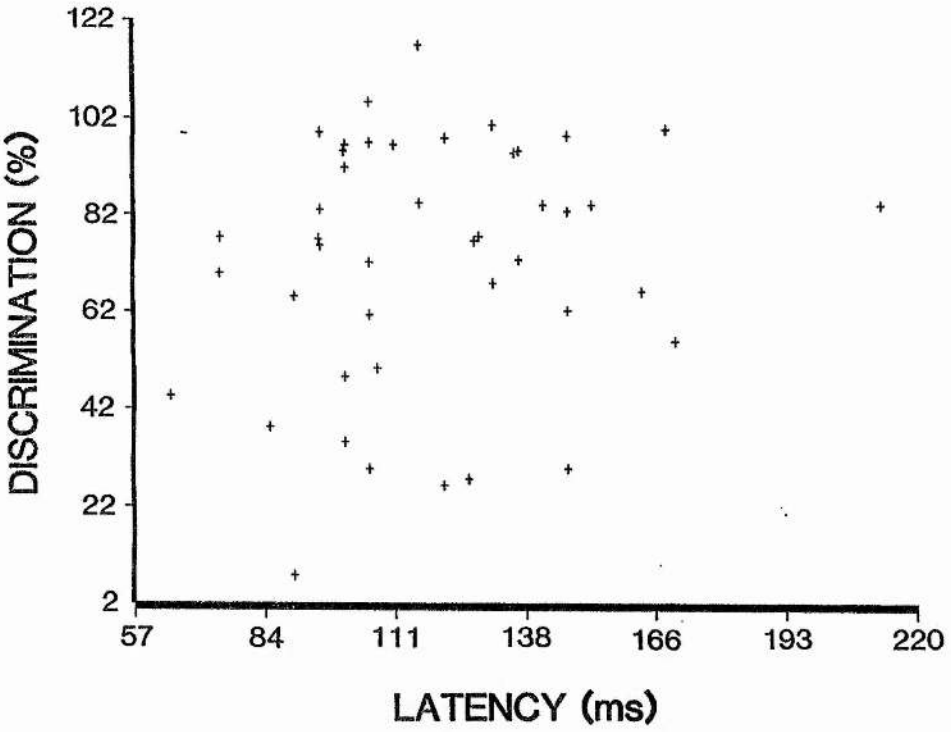


FIGURE 3.7. SCATTER PLOT OF LATENCY AGAINST DISCRIMINATION DURING THE FIRST 100 MS OF RESPONSE. Each cell (n=44) is represented by a single cross. Discrimination is expressed as 100 * (Best response - Worst response) / (Best response - S/A). There was no significant correlation between the two measures ($r_{[42]} = 0.169$, $P > 0.25$).

Snedecor and Cochran 1980] was performed to allow for this relationship, or a Poisson distribution underlying firing rate, which might invalidate ANOVA assessment of discrimination. This showed significance within the first 5 ms of response ($F_{[2,42]} = 5.746$, $P = 0.006$). To control for the mean firing rate being proportional to the standard deviation, ANOVA was also performed on response rates in each time bin using the natural logarithm of firing. Discrimination was still evident within 5.0 ms of response onset, ($F_{[2,42]} = 4.871$, $P = 0.012$).

Relation of response and latency

In deriving the time course of responses for the Average Cell the synchronization of the onset of response of the different cells to the same moment in time could potentially obscure a slow time course of discrimination within a sub-population of cells. Cells with long latency but quickly discriminating responses could, in principle, mask the responses of cells with early response latency but slow discrimination. This potential artifact, however, does not appear to be a problem with the sample of cells studied here because latency was not found to correlate with any of the discrimination parameters (see Table 3.4). The lack of correlation between latency and discrimination confirms that good discrimination of static head views is seen in both cells with early responses and cells with late responses (Figure 3.7).

To examine the relation between response and latency further, the responses of 6 cells with the shortest latency were compared to those of 6 cells with the longest latencies and to 6 cells with mid range latencies. The 6 cells with early latencies exhibited slightly higher peak firing rates in both the Best and Worst category response than cells with mid or long latencies. This is reflected in the negative correlation of peak response magnitude (peak - S/A) with latency (Best category: $r_{[42]} = -0.366$, $P = 0.015$; Worst category: $r_{[18]} = -0.625$, $P = 0.003$) and over the first 100 ms of the response (Best category: $r_{[42]} = -0.411$, $P = 0.006$; Worst category: $r_{[18]} = -0.557$, $P = 0.011$). For each of the 3 sub-

Table 3.4

Parameter Correlations with Latency

Parameter	r	d.f.	p
Timing			
Rise Time (ms)	-0.177	42	0.2511
1/2 Fall Time (ms)	-0.123	42	0.4255
Decay Time (ms)	-0.021	41	0.8937
Duration (ms)	-0.099	41	0.5264
Firing Rates			
S/A (Spikes/s)	-0.272	42	0.0739
Peak (Spikes/s)	-0.360	42	0.0164
1st 100ms (Spikes/s)	-0.409	42	0.0058
2nd 100ms (Spikes/s)	-0.247	42	0.1053
5th 100ms (Spikes/s)	-0.254	42	0.0963
End 100ms (Spikes/s)	-0.208	42	0.1756
Normalized Response Magnitudes			
1st 100ms : Peak (%)	-0.289	42	0.0574
2nd 100ms : Peak (%)	-0.071	42	0.6462
5th 100ms : Peak (%)	-0.033	42	0.8320
End 100ms : Peak (%)	0.011	42	0.9435
Discrimination Measures			
Peak Disc (%)	0.156	42	0.3107
1st 100ms Disc (%)	0.169	42	0.2741
2nd 100ms Disc (%)	-0.112	41	0.4758
5th 100ms Disc (%)	0.076	41	0.6297
End 100ms Disc (%)	0.101	40	0.5226

The correlation coefficient (r), degrees of freedom (d.f.) and probability are listed for each parameter. S/A, spontaneous activity.

populations (short, mid and long latencies) time synchronization of response onsets indicated immediate discrimination between Best and Worst responses within 5 ms of response onset (for the 6 cells with the shortest latency $F_{[1,5]} = 12.78$, $P = 0.016$). Thus while discrimination was evident at the onset of response for all cells, those with short latencies gave larger absolute responses to less effective stimuli.

DISCUSSION

Response properties to effective stimuli

The most striking aspect of the STPa cells' responses was the initial transient burst of activity. After this burst, firing rate decays gradually to approximately 15% of its peak rate. The response magnitude of individual neurons to effective stimuli was large, typically reaching 100 spikes per second above spontaneous activity within 50 ms of response onset. Almost half of this increase in firing rate is reached in the first 5 ms of the response. Indeed the stimulus presentation control techniques may have artificially slowed or smeared the initial rise phase since the stimulus presentation hardware had a rise time of 15 ms. This rapid increase of neural activity was not predicted from PDP models utilising lateral inhibition or feed-back loops.

The very rapid rise suggests that a large number of input synapses to STPa cells are active during this short period. It has recently been suggested that typically 5 or more *near-simultaneously* active synapses are required to generate a single spike in temporal cortex (Gochin et al. 1991). This provides a lower bound on the number of active pre-synaptic cells. Under normal circumstances more might be required, since each synaptic relay in the pathway leading to activation of the recorded cell would tend to decrease the tightness of the timing. Indeed the

mean estimate of the number of pre-synaptic spikes to activate a post-synaptic cell was 39 (Gochin et al. 1991).

The rapid rise of activity following response onset is consistent with the importance of the first or first few spikes to coding (Thorpe and Imbert 1989). Given the exacting time constraints of processing (see later), information must be passed on to the next stage on the basis of this initial activity.

Response differences between Best, Mid and Worst categories

Following the initial transient peak of activity, the rate of decay (as measured by Duration, Half-fall and Decay parameters) was found to vary across the category of response. The decline was progressively faster for less effective stimuli. As the difference in decay rate was found in most cells (38/44, 86%), it suggests that this reflects a general property of the neural responses in temporal cortex.

The between-category differences in response decay could reflect interactions between STPa cells, feed-back loops involving later or earlier processing stages, or interactions between cells in earlier stages. It should be stressed, however, that there was no evidence for inhibitory interactions during the first 80 ms of response.

Temporal resolution of the analysis

The analysis of the neural responses was performed using 4.8, 5.0 and 5.2 ms bin sizes. These allowed separation of single spikes up to frequencies of 208, 200 and 192 spikes per second respectively. The mean peak firing rate of the population was 115 spikes per second and the maximum mean peak firing rate observed in any one neuron was 256 spikes per second. The dynamics of the

neural responses have therefore been analysed at a temporal resolution matched to the spike frequency exhibited by the cells studied.

The temporal resolution also matches that for synaptic transmission. Allowing for EPSP summation where appropriate (see Miles and Wong 1986; Walmsey and Stuklis 1989; Miles 1990), estimated minimum transmission times are, for guinea-pig hippocampal neurons, 2.9 ms (Miles, 1990), 7.3 ms (Sayer et al. 1990) and 12 ms (Miles and Wong 1986). In rat VI, Mason et al. (1991) report a mean of 2.5 ms for transmission between neurons separated by only 140 μm. Transmission between cat VI layers involves lags of at least 5 ms (Best et al. 1986; Maunsell and Gibson 1992). Transmission between V1 and superior colliculus is at least 4.6 ms based on antidromic spike propagation from the primate superior colliculus to V1 (Findlay et al. 1976).

Time estimates depend on the proximity of neurons and the difference between the post-synaptic cell's resting and threshold potentials. It is likely, however, that 5 ms will elapse between detection of an impulse in the soma of an input neuron and impulse generation in a second target neuron in a different brain area. Transfer of activity between cells within the same area might be slightly quicker (say 4 ms) because of the greater cell proximity. The responses and discrimination measures between stimulus patterns has thus been calculated at a time scale that is fine enough to separate individual units of processing (action potentials).

The time course of discrimination

The method adopted to establish a measure of discrimination is similar to the pattern model of Geisler and colleagues (1991) but does not assume an `ideal observer' which extracts all available information (c.f. Geisler et al. 1991): rather the assumption is only that the firing rate of pre-synaptic neurons gives a degree of certainty about the stimulus meaning (Barlow 1972, 1985).

Discrimination in the population response (Figure 3.4b) occurred some 20 ms after the response onset. This delayed discrimination arose because assessment of the population response was made with cells with a variety of onset latencies. As response latency itself did not correlate with the quality of initial discrimination (i.e. both short and long latency cells exhibited good initial discrimination), the dynamics of the Average Cell discrimination was calculated. This showed discrimination between response categories within 5 ms of response onset. With a very large cell population discrimination would take place at the population latency. Indeed, the six cells with a sub-population latency of 70 ms discriminated head views within 5 ms response onset. Thus, discrimination between head views in the STPa occurs as early as 70 ms after stimulus presentation.

The emphasis that discrimination between input patterns is possible within the first 5 ms of the response, this claim does not deny the possibility that the quality (statistical reliability) of discrimination increases over time. Indeed, using 5 ms time windows, a trend for increased significance of discrimination can be seen over the early part of the response of the Average Cell (Figure 3.5b). The peak variance ratio ($F_{[2,42]} = 21.1$, $P < 0.0001$) was obtained some 80 ms after response onset.

It has been proposed that the temporal changes in individual neural responses occurring over hundreds of milliseconds contain information about the stimulus pattern (Richmond et al. 1987; Richmond and Optican 1987; Optican and Richmond 1987; McClurkin et al. 1991a,b; Gawne et al. 1991a,b; Eskandar et al. 1992a,b). Such temporal coding of firing rate is independent of the dynamics of cell responses studied here. Further, the results of Richmond and Optican have recently been questioned by Tovee and Rolls (1993), who found that using a similar, but more appropriate correction for small sample size, most information is coded by neuronal firing rate and relatively little information is coded in the temporal pattern of discharge. The study of Tovee and Rolls (1993) also shows

that most (84%) of the information within the neural firing is available within the first 50 ms of the response, and indeed a substantial amount (44%) is present within the first 20 ms of the response.

Similarly, a number of studies have speculated that temporally coherent oscillations of firing within cell populations may code information about the nature of stimuli (e.g., Gray et al. 1989; Engel et al. 1992; Singer et al. 1990a,b). The role of such oscillations is, however, unclear in primate cortex (Young et al. 1992; Tovee and Rolls 1992; Kreiter and Singer 1992) and still debated (see Young et al. 1992; Singer et al. 1992). This study has not focussed on whether information about stimulus form can be derived from the temporal pattern of response over an extended period of time; rather it has addressed the issue of the latency at which the response of a cell discriminates between image patterns at a statistically reliable level.

The conclusions about the presence of discriminatory activity as soon as the response commences do not contradict claims that additional information is present in the temporal pattern of discharge (although no investigators have suggested how this information could be reliably transmitted to other cells). Temporal codes over long time scales may well have a role in or reflect processes involved in, pattern recognition (e.g. they could reflect top-down guidance of processing). Such slow time course effects, however, do not appear to have a role in the initial discrimination of patterns which is achieved extremely quickly.

Several detailed models of cortical visual processing predict that cells in visual cortex should initially exhibit broad tuning, with selectivity increasing over time (e.g. Krone et al. 1986; von Seelen et al. 1987; Rybak et al. 1991). Indeed, there is evidence for the role of lateral inhibition in tightening the orientation tuning curves of some V1 cells (Sillito 1975a,b; Berman et al. 1991; Hata et al. 1991; Best et al. 1986).

However, recent reports indicate very rapid pattern discrimination (present from response onset) in the visual cortex, even for early latency cells (Trotter et

al. 1989; Vogels and Orban 1991; Thorpe et al. 1989; Douglas et al. 1991; Knierim and Van Essen 1992; Li et al. 1993; Miller et al. 1993). This is perhaps not surprising as many features of simple cells can be accounted for by the afferent input from the LGN (Chapman et al. 1991; Worgotter and Holt 1991; Worgotter and Koch 1991; Reid et al. 1991). The finding that the highly non-linear response properties of many V1 cells can also be predicted from feed-forward networks (Lehky et al. 1992) is also encouraging.

It is unlikely that the rapid discrimination seen in the STPa cells studied here was due either to lateral inhibition or feedback influences. Lateral interactions would require one synaptic relay (4-5 ms) before their effects can be seen (see following chapter for detailed discussion). If discrimination depended on lateral inhibition then it should be weak in the first time bin (5 ms) and improve thereafter. Discrimination was, however, strong within the first 5 ms of response, with no evidence of relative suppression of weaker response until some 80 ms after the response onset (Fig 3). The present analysis does not, however, exclude a role for lateral influences of an excitatory nature. Feed-back loops require a minimum of 2 synaptic relays (9-10 ms). If discrimination depended upon feedback, discrimination would not be evident within 5 ms.

It is possible that lateral connections and feed-back loops located at earlier stages of processing are "thresholded", and this "cleaned up" information passed to STPa cells. With such an operation, one would expect input from effective stimuli to reach the threshold earlier than those to less effective stimuli (Thorpe 1990). This would lead to increased response latency and rise time for poorer responses. Statistically this was not found (Tables 2 and 3b).

Thus, rapid pattern discrimination at the neural level may be evident at early stages of visual analysis in striate cortex and, as shown in this study, is apparent in the temporal cortex. These empirical observations lie in contrast to the performance of many models employing artificial neural elements (see chapter 2), though some do show rapid discrimination between input patterns, especially

when operation relies on feed-forward connections (e.g. Foldiak 1990; Grossberg 1987).

Role of attention in encoding stimulus properties

It could be argued that the rapidity of cell response (< 90 ms) and discrimination between complex patterns (< 5 ms) observed were due to the experimental subject `learning to expect' a small number of testing stimuli and paying preferential attention to relatively simple visual cues. The mechanisms by which this could be implemented are not clear but might involve `pre-facilitation' or `pre-inhibition' to enhance sensitivity to the relevant cues.

There are two points to be made with respect to this issue. First, a very large set of stimuli (> 300) was used for examining responses to visual information about the head and body. Each block of trials consisted of head views, control and blank stimuli presented in random order. Within each block, control stimuli were randomly selected from a large set (>100) of different views of different objects. Thus, there were no predictable visual cues that could be `pre-selected'.

Second, the subjects performed a behavioural task (LED colour discrimination) during presentation of test visual stimuli. Performance in this task was > 80% correct, with reaction times less than 500 ms. The subject could obtain multiple rewards (licks) only if the discrimination of LED colour was performed at these fast reaction times. Objective measures of eye position showed that fixation patterns could not account for differences in neural responses to different stimuli (Figure 3.2). Eye movements by themselves are not *per se* a good indicator of attention and the head view stimuli were not used as disciminanda. In the behavioural task used, however, the subject had no restrictions on eye movements. Thus it is likely that the fixation point at any time was a good indicator of the subject's focus of attention. It seems reasonable to assume that any

`pre-selection' of visual cues was therefore directed more to the colour of the LED than to simple visual cues related to the test stimuli, since the test stimuli were themselves irrelevant to task performance.

The use of a warning tone just prior to stimulus presentation can suppress cell pre-stimulus activity of cells in IT cortex (Sato et al. 1980). Relatively minor effects on the amplitude of STPa and IT cell responses have been observed from paradigms using a second concurrent visual stimulus or fixation spot; these include both decreases (Richmond et al. 1983) and increases in response rate (Moran and Desimone 1985; Richmond and Sato 1987; Sato 1988; Fuster 1990). As these studies did not report changes in response latency, it seems likely that attention in the behavioural task, which was to an LED and not the test stimuli, may have caused a slight decrease in response magnitude to the test stimuli but no other effects.

Pattern ambiguity and rate of discrimination

Models using feed-back loops exhibit particularly slow discrimination when two input patterns are ambiguous. It is relevant therefore to consider the extent to which different head views constitute similar input patterns. The cells in the present analysis show a broad tuning curve for perspective view (half width, half height = 60 degrees) and have a minimum response some 90-180 degrees from the optimal head view (Perrett et al. 1991). Given this breadth of tuning, stimuli producing Best, Mid and even the Worst levels of responses would have many simple features in common (e.g. those associated with eyes, hair and mouth). That the patterns share features is indicated by the observation that there was a small response in the Worst category. Discrimination between views is therefore a situation where recognition models would tend to show slow discrimination because the input patterns are similar.

Work by Tanaka et al. (1991) shows cells in the inferotemporal cortex (IT) which require specific combinations of features, not only to be present but also to be in particular spatial relationships. A more direct demonstration of selectivity to configuration of STPa cells sensitive to the face has been shown by Perrett et al. (1982). Recently, Miyashita has reported cell selectivity within IT indicating that the effect of stimulus rotation/reflection seems to have less of an effect of response magnitudes that changing the relative spatial arrangement of the component features (Miyashita 1993). With spatial selectivity as well as simple feature content sparsely coded, output from such IT cells could be used to establish the discrimination between views of the head that found in STPa cells.

Number of processing steps prior to STPa activation

Visual responses of cells in V1 have a mean latency of 53 ms (Robinson and Rugg 1988), with a range of 40-100 ms (Vogels and Orban 1991). Shorter latencies are possible in layer IV with stimuli presented at high light intensity, contrast and fast rise time (Maunsell and Gibson 1992). Taking 35 ms as the earliest latency estimate for V1 cells with the stimuli used, this leaves 35 ms of time available for processing between V1 and the initial activation of some STPa cells (70 ms).

While the shortest route from V1 to the anterior STPa may be via V5 (MT) this route is generally regarded as a `motion' pathway (Zeki and Shipp 1988). The transmission of `form' information is more likely to involve a a ventral route involving visual areas V1, V2, V4, TEO and TE (Ungerleider and Mishkin 1982; Harries and Perrett 1991; see also first chapter). This route could involve a minimum of 3 areas (missing out V2, Felleman and Van Essen 1991). Recent work of Baizer et al. (1991) suggests a relay through a further area (e.g. TEa) would be needed to reach the site of recording (STPa upper bank, areas TPO and

PGa, Harries and Perrett 1991; Perrett et al. 1991) since inputs from TE and TEO are almost exclusively to the lower bank and fundus of the STPa.

Feed-forward inputs to cortical areas enter layer IV, whereas outputs from an area predominantly exit from cells in the infra-granular layer (VI) or supra-granular layers (I-III) (Felleman and Van Essen 1991). Thus transmission of form information between V1 and the recorded cells in STPa probably involves a minimum of 8 synapses (through 4 visual areas, each with an input and output layer). Coupled with the transmission times suggested above this would give the earliest latencies of activation (ms) by areas as follows:-

V1 35; V2 44; V4 44-53; TEO 53-62; TE 53-71; TEa 62-80; STPa (TPO/PGa) 71-89

(Using the nomenclature of Felleman and Van Essen the temporal areas would be PIT, CIT, AIT and STPa)

These considerations should make it clear that to set up the first one or two spikes in early latency STPa cells, the flow of information has to be entirely feed-forward.

Generality of results to processing in temporal cortex

The neural responses examined here all showed evidence of a transient burst (for 100-200 ms) followed by a lower more maintained discharge lasting up to 800 ms. This pattern of activity is apparent in illustrated examples of responses of temporal cortex cells from many previous studies (e.g. Richmond et al. 1983; Fuster 1990; Perrett et al. 1985; Sato 1988, 1989). The estimates of spontaneous activity from the present study (mean of 8.6 spikes per second) also seem typical of IT and STPa neurons reported previously (Ridley et al. 1977; Mistlin and Perrett 1990; Perrett et al. 1991). The high peak response magnitudes (200 spikes/s) observed here are not uncommon in other studies (e.g. Richmond et al. 1987). Previous estimates of the range of response latencies in IT and STPa cortex also suggest that the cells studied here are typical (mean 142 ms, Ridley et

al. 1977; 144 ms, Fuster 1990; range 70 to 220 ms, Richmond et al. 1983; 73 to 120 ms, Richmond et al. 1987; 52 to 400 ms, Coburn et al. 1990).

The similarity of the response properties of the sub-population of cells studied here to the properties of cells from other studies of the temporal cortex suggests that the conclusions about the dynamics of discrimination between different head views are probably applicable to the processing and discrimination of a variety of complex visual patterns. It is common for published records of visual cells in IT (and elsewhere) to display responses to effective stimuli which show a large, clear onset. These records also show, for the same cells, no response to non-effective stimuli. As with this study, this response pattern implies discrimination at the onset latency.

Coding of information

Within 5 ms of information arriving at the STPa cortex, the output from about 150 neurons would contain sufficient information to distinguish reliably between head views separated by some 60 degrees of rotation. (the estimate of the responses of the Average Cell was based on data from 22 cells with 5-10 trials in each response category for each cell). This provides an indication of the number of cells necessary to achieve discrimination between head views *within 5 ms*. Each cell individually showed discrimination between views when the responses were assessed *over a 250 ms period* for 5 trials of each response category.

The present study is consistent with others in indicating that sparse coding using a small number (e.g. 20, Young and Yamane 1992) of cells in the temporal cortex can achieve an effective level of discrimination between complex stimuli. The findings here indicate that by utilizing a greater number of cells (150), discrimination can be achieved in a very short time scale (5 ms). Temporal integration by post-synaptic cells receiving the output from the population studied here would lead to greater discrimination between the stimuli. This is probably of

particular importance some 400 ms after stimulus onset and thereafter, where discrimination is not statistically reliable when sampling over a 5 ms period, though longer sample periods (e.g. 100 ms) the discrimination is statistically still reliable (Table 2).

The assessment of the discrimination was performed using 2-way ANOVA, with cells as one factor. On the average, there is a statistically reliable difference between responses of the three categories. There are however differences in the absolute firing rates of the individual neurons. It is an implicit assumption in such an analysis that, for the brain to utilise the differences between response level, a mechanism must exist which can accommodate any differences in absolute firing rates of the individual input cells. Differential synaptic efficacies, or synapse location at different electrotonic distances are two possible mechanisms.

The results indicate that at least three levels of response rate are reliably differentiated. This shows that STPa cell response levels were not just binary (effective stimuli present or absent) but were graded in a quantitative manner. Such graded discrimination between input patterns can be seen as providing `evidence' proportional to the confidence that a given stimulus pattern is present in the image (Barlow 1972, 1985). Again it should be stressed that it was possible to grade the information relayed by the sub-population of cells, not just at the peak of cell firing rate but also within 5 ms of cell response onset.

The rapidity of discrimination (within 5 ms) and the firing rate (rarely in excess of 1 spike in a 5 ms period) combined with the fact that any one input is unlikely to have sufficient synaptic strength to drive a post-synaptic cell by itself within 5 ms (Gochin et al. 1991) suggests *it is the presence of single spikes in multiple input sources* (rather than the number of spikes from one input) which is necessary for the discrimination seen between input patterns. This could be paraphrased as "Grandmother cell coding could only work if such cells had a grand 'mother' of a synapse to the next cell". Even with strong synaptic

connections, single cell coding could not account for the rapidity of the graded discrimination observed, as a single transmitting source could only differentiate between patterns in a binary fashion within this time (1 impulse). This in turn implies that multiple sources are required to transmit the graded information to the next stage within 5 ms,. Thus, in contrast to Barlow's doctrine (Barlow 1972, 1985) consideration of the time constraints of coding suggests that the information about stimulus certainty requires the pooling of activity across a small population of cells (each with similar stimulus selectivity).

The `dual-coding' principle recently proposed by Kruger and Becker (1991) suggests that coding of information is by both the precise time at which impulses occur and their presence or absence. Whilst the data support the notion that it is the presence of impulses that codes the information within one cell, the interpretation offered here stresses that it is the number of single impulses seen within any one short time period *across the input population* that codes information about the image. This interpretation, that it is input synchrony has recently been proposed by Singer (1993).

It has been argued that a coding system based on spike arrival times can give very rapid discrimination (Thorpe 1990): the first cell to transmit information to the next level becomes a `winner-takes-all' in a purely feed-forward manner by using an inhibitory `veto' of late spike arrivals from competing inputs. With such a mechanism and sampling across an input population with redundant coding, processing of the image could be accomplished with great rapidity in a feed-forward way, allowing the rapid discrimination described in this study. This would not be possible with a coding system relying on many spikes from just a few neurons.

Physiological studies on the issue of population vs single cell coding are often made by comparing the quality of discrimination at the single cell and behavioural level (for discussion see Barlow 1985). Even if the stimulus selectivity of individual cells is extremely high (and equal to the psychophysical

observer) multiple cells with the same high selectivity would be required to transmit the information through the nervous system at the required speed. This speed is reflected both in the rapidity of behavioural responses and in the rapidity of discrimination between neural responses reported here.

Transmission and representation of information how many cells are needed?

The findings from this study indicate that rapid initial discrimination is maintained throughout the visual system. What is more, the speed of information flow and the very fast rise in firing rates combined with the absolute levels of firing rates of cortical neurons suggests that a multiple redundancy coding system is used with information being passed from one processing area to another on the basis of the first (and possibly second) spike relating to the stimulating input. The analysis also implied that multiple sources are required to transmit the graded information to the next area within 5 ms. This raises the question of how many cells are required.

Barlow (1994) suggested that coding might require several 'cardinal cells' rather than a single 'pontifical' or 'grandmother' cell. Similarly Konorski (1967) noted that coding might include several 'gnostic' units for important objects. Consideration of the time constraints of coding suggests that the information transmission requires the pooling of activity across a small population of approximately 150 cells. The notion of 'cardinal cells' or 'gnostic units' can be supported from these studies but only if there is a certain degree of redundancy, with approximately 150 cells showing nearly identical response selectivity and response latency.

The coding of information using multiple cells with similar selectivity has a potential problem: if the inputs to all the cells are correlated (or more strictly, if the noise between cells were correlated) then improvement in signal to noise from utilising many units would be limited (i.e. a greater number of inputs above an

upper bound would not improve the quality of pattern discrimination). Groups of approximately 10^5 cells with similar stimulus pattern selectivities are found in a group or column in IT and STPa cortices (Fujita et al. 1992; Perrett et al. 1984, 1986). The work of Gochin et al. (1991) and Gawne and Richmond (1993), however, has shown that even cells with similar stimulus selectivities have only partially correlated noise during their responses. Even with this small amount of correlated noise Gawne and Richmond (1993) have argued that it is unlikely that coding of stimulus attributes in IT is distributed over more than a few (20) units. This is consistent with studies which have shown a sparse population of IT cells can code accurately detailed pattern (face identity) information when assessed over a long (200-1000 ms) time period (Young and Yamane 1992).

The small number of units that can be used to code object representation efficiently (20) is somewhat smaller than the numbers (at least 100-150) argued here as necessary to explain the rapid discrimination seen between input patterns. There are two points which are relevant to this apparent discrepancy. First, the work of Tanaka and colleagues (Fujita et al. 1992) indicates that there may be several modules or columns processing similar information. The distances between these columns would make it very unlikely for them to share common input (i.e. have correlated noise, Gochin et al. 1991). Therefore it is reasonable to assume that each column involved in processing information about one particular feature could pass this information forwards using 20 or so neurons to achieve optimal signal to noise ratio. Only five to ten columns would then be necessary to achieve the appropriate total number of cells. Second, it is argued for multiple lines of transmission using only one spike on each line to enable *rapid* discrimination at the subsequent level: in this situation the effect is not one of improving signal to noise but rather of reducing the transmission time of the information (at a given signal to noise ratio). Thus it seems that a small number of cells may be sufficient to represent particular aspects of an object adequately (and indeed may reflect the highest degree of accuracy of representation possible),

whereas the observed speed of information transmission requires multiple cells as inputs to those representations.

CHAPTER 4

PROPERTIES OF PDP MODELS PREDICTIONS OF PATTERN DISCRIMINATION DYNAMICS FROM A SIMPLE SIMULATION

(see Oram & Perrett 1994, *Neural Networks*, 7, 945-972,1994)

INTRODUCTION

In this chapter it is argued that most network or parallel distributed processing (PDP) models processing visual patterns, or indeed any complex pattern of inputs, suggest a gradual emergence of the discrimination between visual patterns. The dynamics of PDP models are rarely compared with neural responses, yet, as seen in the previous chapter, the macaque ventral pathway appears to process information in a primarily feed-forward manner. It is argued here that the slow discriminatory properties of many PDP models, even after extended learning periods, are a consequence of the architecture of the connections between processing elements (nodes) rather than an effect of the processing elements themselves.

Lateral (feed sideways) inhibition is frequently used to apply `contrast enhancement' (Grossberg 1987) between two or more nodes that are both initially activated to similar levels by one input. An effect of such lateral inhibitory connections is a relatively slow rising phase of activation. This slow rise is characteristic even of the eventual `winning' node, since initially it is itself being inhibited by other competing nodes. The precise rate of activity increase will of course depend on the time course and strength of inhibitory connections.

The presence of `re-entrant' feed-back (Edelman 1978) from higher network layers to lower layers can also contribute to a gradual emergence of discrimination between input patterns. Feed-back loops from initiating nodes in a

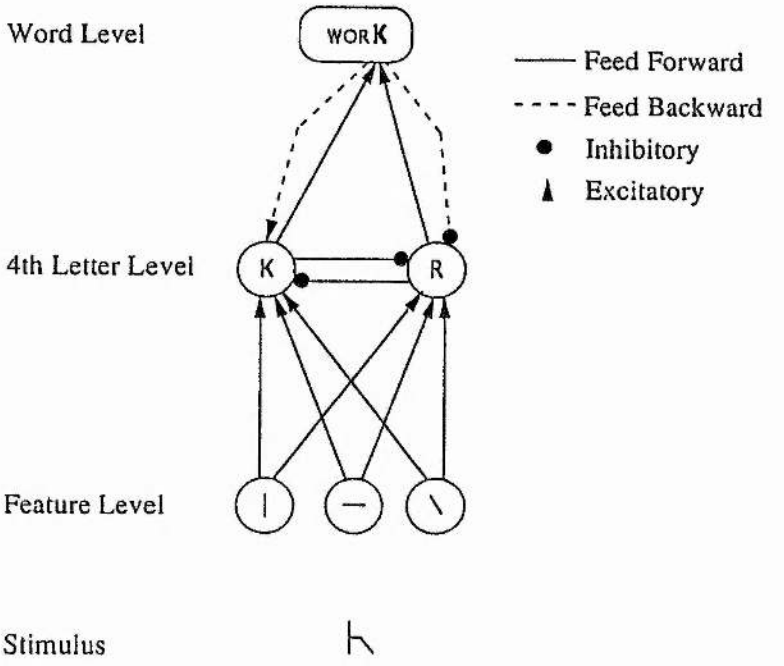


FIGURE 4.1. A MODEL ARCHITECTURE WITH LATERAL INHIBITION AND FEEDBACK. The representation of part of a model for 4 letter word recognition with lateral inhibition and feedback connections. For details see text.

higher layer act on nodes of a lower layer which provide *inputs* to the higher layer. Inputs to competing nodes in the higher layer can be suppressed through inhibitory feed-back connections and, simultaneously, positive inputs to the initiating nodes can be supported through excitatory feed-back connections. It is important to make a distinction between feed-back connections (top-down) and models using back-propagation to modify synaptic weights, which most investigators do not regard as biologically feasible (Grossberg 1987).

Both feed-back and lateral connections are particularly useful in enhancing discrimination between input patterns which are impoverished or noisy but such connections can force the discrimination between input patterns to emerge gradually. The slow temporal characteristics are found in many models especially when distorted or noisy input is used (e.g. Frohn et al. 1987; Rumelhart and Zipser 1986; McClelland et al. 1986; Seibert and Waxman 1989, 1990, 1992a,b).

To clarify the reasons behind this claim it is helpful to provide an example. Figure 4.1 shows a section of a model from McClelland et al. (1986) for 4 letter word recognition. The section shown represents only a small part of the complete model. The basic features of the model consist of a set of feature detectors (oriented lines of given length in given positions. The outputs of these feature detectors could be combined to form any single letter. As there only four letter words used, the structural pattern of feature detectors feeding onto letter detectors was repeated for each of the four possible letter positions. These letter detectors in turn fed into word detectors. Thus under normal circumstances the model would combine activation of the feature detectors to activate the letter detectors and these in turn would combine to activate the word detectors. Possible ambiguity between word detector activation levels is "contrast enhanced" by lateral inhibition between (a) the word nodes and (b) the letters and also (c) by the use of appropriate feedback excitation and inhibition connections (see figure 4.1).

If the word "WORK" was presented to the model there would be no difficulty in the 'WORK' node of the model reaching full activation, but if the

final letter position stimulus was degraded as indicated in figure 4.1, then the model would behave differently. When the stimulus is first presented, the appropriate feature detectors become active. A finite time later the detectors in position 4 of letter level would become active, but in this case both the 'K' and 'R' nodes would be equally excited. Another finite time later, whilst inhibiting any other letter detectors in the fourth position, the 'K' and 'R' letter nodes also inhibit each other to an identical extent. At the same time that this lateral inhibition occurs, the letter feature activations are passed forward to the word detector level. It is during the next processing cycle that the word to letter feedback connections start to come into play. As there is a word "work" but not "worr" (and the 1st to 3rd three letter positions as well as the 'K' node are activating the 'WORK' word node) the 'K' node receives feedback excitation from the 'work' node. At the same time the 'R' node receives feedback inhibition from the 'WORK' word node. It is only at this stage that there is a difference between the total input activation levels of the 'K' and 'R' nodes. The activity difference this differential input produces subsequently will be increased (enhanced) by both continued differential feedback input from the word level and (because of the influence of the differential top-down influence from the word level nodes) also by direct lateral inhibition between the two letter nodes.

In this extreme case the activity of the two letter detectors only becomes separated after the feedback connections have played their role. This effect can only occur two computational cycles after the initial activation of the letter detectors (1 cycle to pass the information forward to the word level, 1 to be passed back down). In a similar way, the effects of lateral inhibition can only be seen one cycle after the differentiation of the activation levels of the 'K' and 'R' nodes (a total of three computational cycles after initial activation).

Thus many PDP or network models of pattern processing exhibit a number of common operating characteristics that are a result of the architecture used. They predict that the time course of response in relevant elements will show (1) a

slow rise to peak activation in the node showing highest activation and (2) once an activity peak has been reached, the network will stabilize and the activation level in the nodes will be maintained until the input pattern changes. These features of slow rise and decay will depend on parameters within the network, including the exact strength and timing of connections. A more interesting prediction is (3) the slow emergence of discrimination at output levels for closely related input patterns.

These predictions are independent of the hardware used as calculation of activation levels in nodes is typically performed at discrete intervals of time (computational cycles). Thus the temporal characteristics will remain qualitatively identical even with faster hardware since the number of cycles needed for the computation will not change, even though the absolute time taken will decrease.

The work described in this chapter was undertaken to investigate the validity of the arguments and predictions outlined above. The simulations also served as an to attempt to generate model data that could be compared to neurophysiological data.

METHODS

The predictions were tested using a simple implementation of an interactive activation competition (IAC) model described by McClelland and Rummelhart (1988). This model was chosen since there are published values available against which the implementation could be checked, thus ensuring that the program was performing correctly. Further, because the connections and weights between the computational elements in IAC models are predetermined and fixed, this class of PDP model has no learning stage and is deterministic.

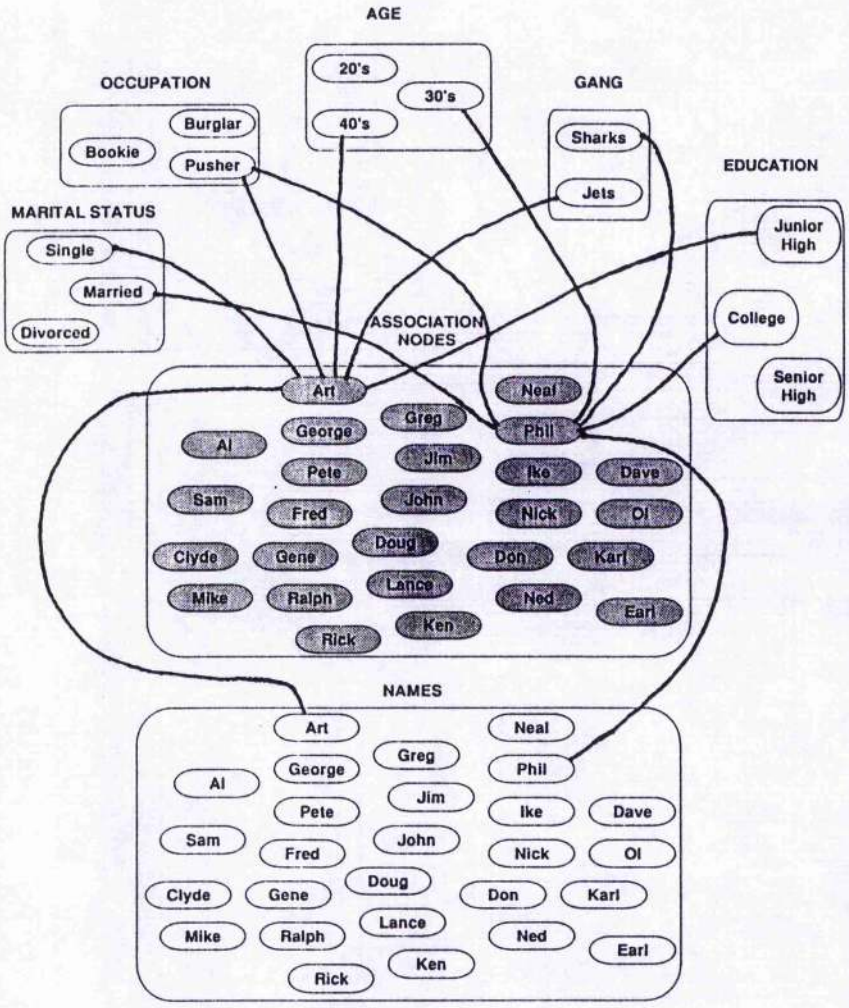


FIGURE 4.2. SCHEMATIC REPRESENTATION OF THE "JETS AND SHARKS" NETWORK. The network used for the simulations was an IAC network. The nodes of the network are depicted by ovals. All nodes within one of the rectangles are connected to all other nodes within the same rectangle by mutually inhibitory connections. All connections between nodes in different rectangles are excitatory. The association nodes (grey ovals) can be used to define the links between names, occupations etc. For example, the 'Art' association node links the name node 'Art' with the 'Single', 'Pusher', '40's', 'Jets' and 'Junior High' nodes. Similarly the 'Phil' association node links the name node 'Phil' with the 'Married', 'Pusher', '30's', 'Sharks' and 'College' nodes. In this way, every name node is connected to a single marital status, occupation, age, gang and education node (not shown).

TABLE 4.1

Network parameters used in the simulations

Maximum possible activation level	1.0
Minimum possible activation level	-0.2
Resting activation level	-0.1
Threshold activation level	0.0
Decay fraction	0.1
Excitatory connection weight	0.1
Inhibitory connection weight	0.1
External input factor	0.4
Input strength	1.0

The network parameters were taken from McClelland et al. 1988. The performance of the network showed identical performance to published figures for given inputs in McClelland et al. 1988.

Network simulations were performed using an implementation of the "Jets and Sharks" interactive activation (IAC) network of McClelland and Rummelhart (1988, and see Figure 4.2). The "Jets" and "Sharks" are two fictitious New York gangs, each member having various properties (name, gang membership, age, occupation, education level and marital status). The basic architecture of this IAC model is one of "pools" of nodes representing properties (e.g. age groups) connected to other "pools". Within each pool, the nodes are fully connected (i.e. each node is connected to every other node) with inhibitory links. Between "pools" the nodes are selectively connected with excitatory links. Finally there is a central pool which acts as a relationship or association center. Conceptually this

gives rise to a situation where within "pools" the nodes are mutually exclusive (i.e. as Phil is in his 30's then he cannot be in his 20's or 40's), with the central pool acting as the link connecting Phils's name with the 20's, Pushers, Married, Sharks and College nodes.

The program was written in Turbo Pascal 5.0 (Borland), using MetaWindows 3.5 graphics routines (Metagraphics Software Corporation), and run on a PC-AT compatible computer (Dell 450DE). Connection strengths, connection factors, minimum maximum and resting activation levels, and external input strengths were all set at the default values (see p. 40 McClelland and Rummelhart 1988, also table 4.1). Their standard update rule for node activity was used. The full list of nodes implemented is given in table 4.2.

Three additional features were added to the standard program. The first was the inclusion of a graphical display mode, whereby the activation levels of the nodes could be displayed over time (computational cycles) to the screen. Screen dumps to encapsulated postscript (EPS) files or direct output to a postscript printer were available. Secondly, random noise could be introduced to each of the nodes in the network. The noise could be simulated as coming from a Gaussian (normal) or linear (uniform) distribution. The level of noise (standard deviation for Gaussian or peak for linear) was expressed as a percentage of the difference between the minimum and maximum possible activation levels of the nodes. The noise was added at each computational cycle to each processing element. The purpose of adding noise was to allow comparison between PDP model nodes and biological data from neurons (see next chapter). Since neurons show random fluctuations in their activity levels and this noise cannot be removed, it was necessary to introduce noise into the model. Finally, when noise was added to the

TABLE 4.2

Characteristics of members of the Jets and Sharks gangs

Gang	Education	Name	Age	Occupation	Marital Status
Jets	Junior High	Art	40's	Pusher	Single
Jets	Junior High	Al	30's	Burglar	Married
Jets	Junior High	Clyde	40's	Bookie	Single
Jets	Junior High	Mike	30's	Bookie	Single
Jets	Junior High	Jim	20's	Burglar	Divorced
Jets	Junior High	John	20's	Burglar	Married
Jets	Junior High	Lance	20's	Burglar	Married
Jets	Junior High	George	20's	Burglar	Divorced
Jets	Junior High	Ralph	30's	Pusher	Single
Jets	High School	Greg	20's	Pusher	Married
Jets	High School	Doug	30's	Bookie	Single
Jets	High School	Pete	20's	Bookie	Single
Jets	High School	Fred	20's	Pusher	Single
Jets	College	Sam	20's	Bookie	Single
Jets	College	Gene	20's	Pusher	Single
Sharks	Junior High	Ike	30's	Bookie	Single
Sharks	High School	Nick	30's	Pusher	Single
Sharks	High School	Karl	40's	Bookie	Married
Sharks	High School	Ken	20's	Burglar	Single
Sharks	High School	Earl	40's	Burglar	Married
Sharks	High School	Rick	30's	Burglar	Divorced
Sharks	High School	Neal	30's	Bookie	Single
Sharks	High School	Dave	30's	Pusher	Divorced
Sharks	College	Phil	30's	Pusher	Married
Sharks	College	Don	30's	Burglar	Married
Sharks	College	Ned	30's	Bookie	Married
Sharks	College	Ol	30's	Pusher	Married

Full listing of the non-association nodes of the simulated model. Each line of the table has an association node which links the non-association nodes (see also Figure 4.2).

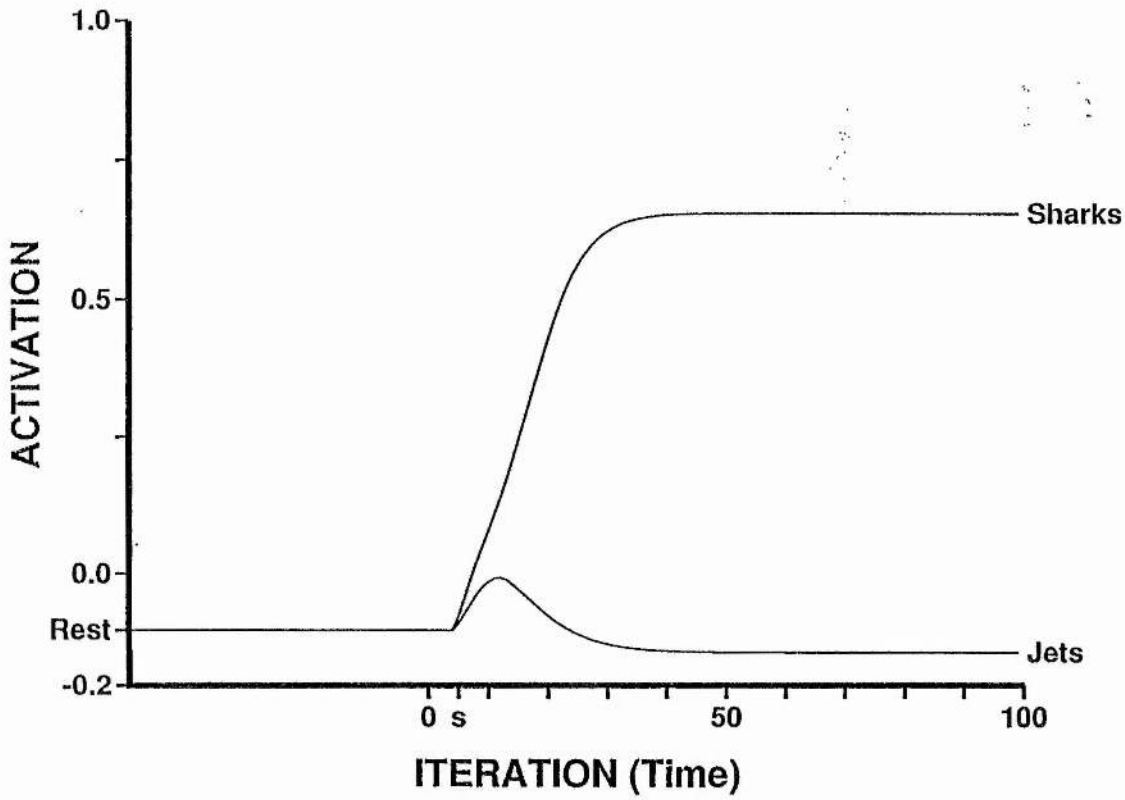


FIGURE 4.3. RESPONSE OF THE NOISELESS MODEL. The activation level against iteration number of the 'Sharks' and 'Jets' node when the 'College' node is set to an activation level of 1. The input starts at iteration 0. The activity levels of the two monitored nodes first separate at iteration 5 (marked 's').

system repeated simulations could be run and the activation values of selected nodes could be stored to disk file at each cycle. After a prescribed number of repeated simulations were completed, the average values of the monitored node activation levels were compared using 2-way ANOVA (node activity as fixed factor, simulation repeat as random factor). The F value for between the average activity of the nodes was taken as the discrimination measure, with a high F value corresponding to successful discrimination between inputs.

The simulations were run with the "College" node set to value one as the input (the highest possible activation level for a node) and the "Jet" and "Shark" nodes were monitored as the output. [A greater proportion of the "Shark" members than the "Jet" members had college education.] The simulations ran for 50 iterations before the "input" was initiated. This was considered analogous to the pre-stimulus period used in the neurophysiological experiments. The initiation of input into the network was taken as the equivalent to stimulus onset. The simulations were stopped 100 cycles after the onset of input.

RESULTS

The performance of the noiseless model

The first experiment was to run the network without the introduction of noise to determine the networks base-line performance. The result of this simulation is shown in Figure 4.3.

As can be seen the network shows that, initially when the monitored nodes first rise above the resting level, the difference between their activation levels is low. As the number of computational cycles increases the difference in activation levels increases. The activation of the 'Sharks' node once above the resting level starts slightly above the activation of the 'Jets' node, and continues to rise until it

asymptotes to a fixed high level. By contrast the activation level of the 'Jets' node starts to rise but then shows a slowing of the rise and finally a decline to asymptote to its final activation level. This slowing of the rise follows the activation of the 'Sharks' node exceeding threshold. Note that the model will always perform in this way and always give the same activation levels (statistical analysis is therefore unnecessary). Further it is possible at any stage after the fourth cycle to determine which of the nodes will remain higher than the other - if one node has a higher activation level than the other it will remain so as long as the input does not change. The tick marked s along the X-axis marks the point where the activation level of the two nodes first becomes different (cycle number 5).

The effect of adding noise

The simulation above was performed without noise and as described suggests perfect discrimination is present whenever two nodes have different activation levels. However, there is clearly an increase over time in the difference of the activation levels which is suggestive that discrimination of the model would emerge slowly if implemented by an intrinsically noisy system (such as the brain).

To test the suggestion that discrimination of the model is slow to emerge, noise was added to the network. There is an important difference between the nature of the noise used here and that used in other studies: previous studies have only added noise to the inputs but have assumed that the calculations performed within the network remain noise free (Frohn et al. 1987; McClelland et al. 1986), although there are a few notable exceptions where noise was added internally to a model (and in some cases found to be an essential feature of the model, e.g. Clothiaux et al. 1991). By adding noise to the computational stages of the network it was hoped to gain insight into the stability of the performance of the network itself and not simply its ability to perform with partial or noisy input. |Simulations using noisy input have also been performed and showed that the network

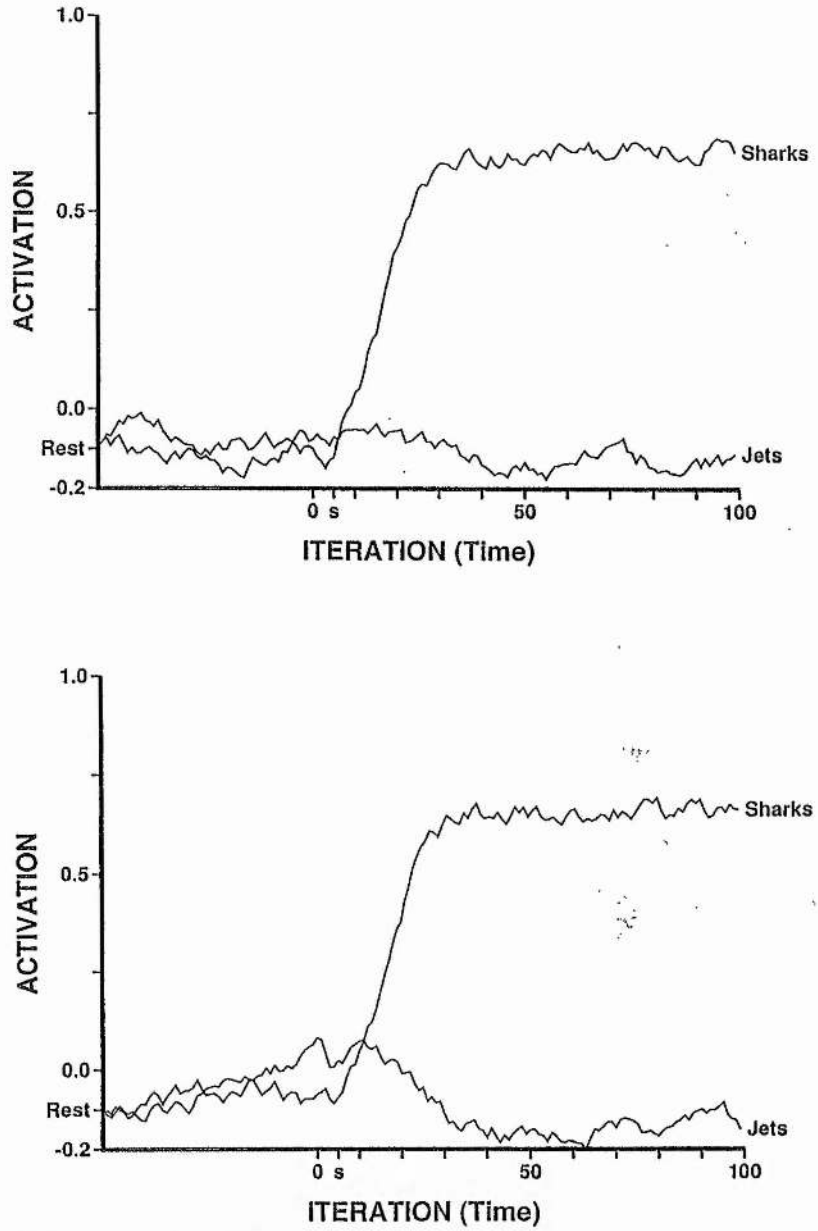


FIGURE 4.4. RESPONSE OF THE MODEL WITH 2% GAUSSIAN NOISE. Two examples of the activation levels of the 'Jets' and 'Sharks' nodes of the network when there is 2% Gaussian distributed noise added internally to the model. The activation level of the 'College' node is set to 1 at iteration 0. The tick marked 's' indicates the iteration where in the noiseless simulation the activation of the 'Jets' and 'Sharks' nodes first separate.

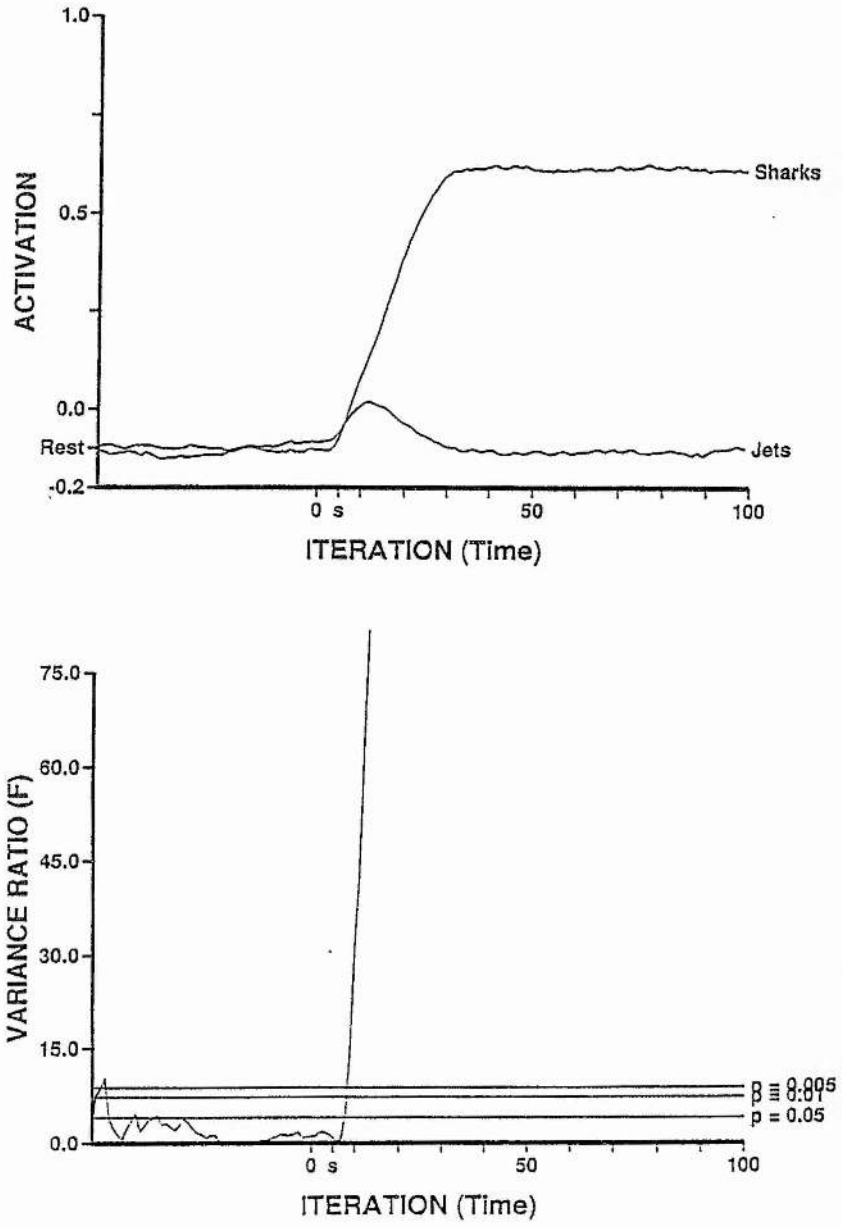


FIGURE 4.5. AVERAGE RESPONSE AND DISCRIMINATION WITH 2% GAUSSIAN NOISE. **UPPER: AVERAGE ACTIVATION LEVELS OF THE 'JETS' AND 'SHARKS'.** The average activation levels of the 'Jets' and 'Sharks' nodes of 20 repeated runs when 2% Gaussian distributed noise is added internally to the model. The activation level of the 'College' node is set to 1 at iteration 0. The tick marked 's' is the iteration where in the noiseless simulation the activation of the 'Jets' and 'Sharks' nodes first separate. **LOWER: DISCRIMINATION.** The F value between activation levels of 'Jets' and 'Sharks' nodes in the upper section. The F values reaches values in excess of 15,000 by the twentyth iteration (not shown). The horizontal lines indicate probability values of 0.05, 0.01 and 0.005. The tick marked 's' indicates the iteration where in the noiseless simulation the activation of the 'Jets' and 'Sharks' nodes first separate.

implemented here is indeed insensitive to noisy input when there is no internal noise.]

Figure 4.4 shows two example trials of the same simulation as shown in Figure 4.3, but with the addition of Gaussian noise added to each element (node) in the network at each computational cycle. The distribution from which the noise was taken had mean 0 and standard deviation of 2% of the difference between the minimum activation level (-0.2) and the maximum possible activation level (1.0). As can be seen clearly this has the effect of removing the reliability of discrimination. Unlike the previous simulation, with noise added it is not the case that if one node is at a higher activation level than the other node it will remain so. The activity levels of the two nodes do, however, show the trend of slowly reaching a quasi-stable level.

The simulation with 2% Gaussian noise was repeated 20 times and the activation levels of the nodes at each computational cycle were stored and then analysed using 2-way ANOVA. The upper section in figure 4.5 shows the average of the activation levels for the 'Jets' and 'Sharks' nodes. As can be seen, the activation levels remain similar for several computational cycles longer than the time when the noise free simulation first showed differential activation levels. Indeed the activation level of the 'Sharks' node only rises above the activation level of the 'Jets' node 2 computational cycles after the noise free simulation first showed a difference in activation levels (marked s in figure 4.5). After this "unsure" or "non-discriminative" period, the average activation levels then separate and reach relatively stable levels. The final activation levels obtained are very similar to those in the noise free simulation. The lower section of figure 4.5 shows the F-value plotted against time (number of computational cycles). It is clear from this graph that statistically reliable discrimination between the activation levels of the nodes does not emerge until several (eight) cycles after the activation levels change from the resting level in the noise free network (marked s). These results were repeated using both another random twenty and fifty

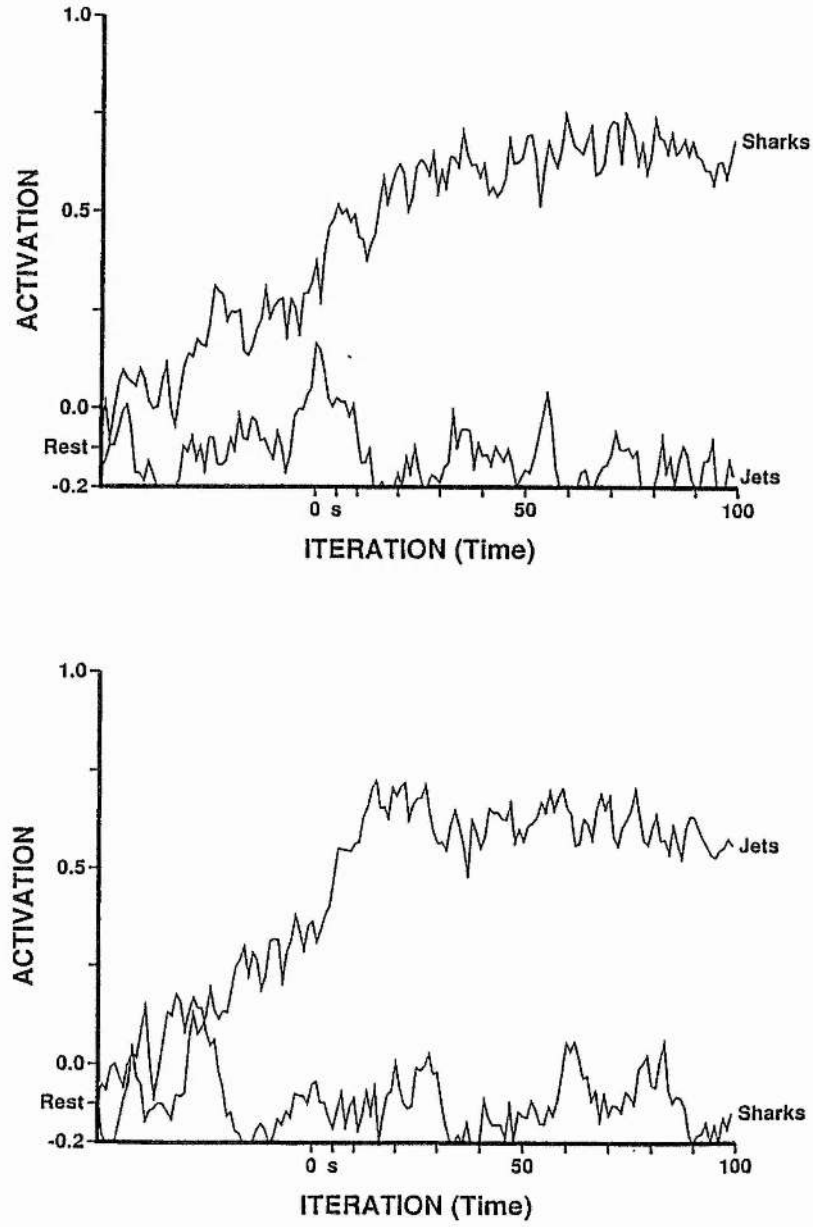


FIGURE 4.6. RESPONSE OF THE MODEL WITH 5% GAUSSIAN NOISE. Same as figure 4.4 except that 5% Gaussian noise was used in the simulations.

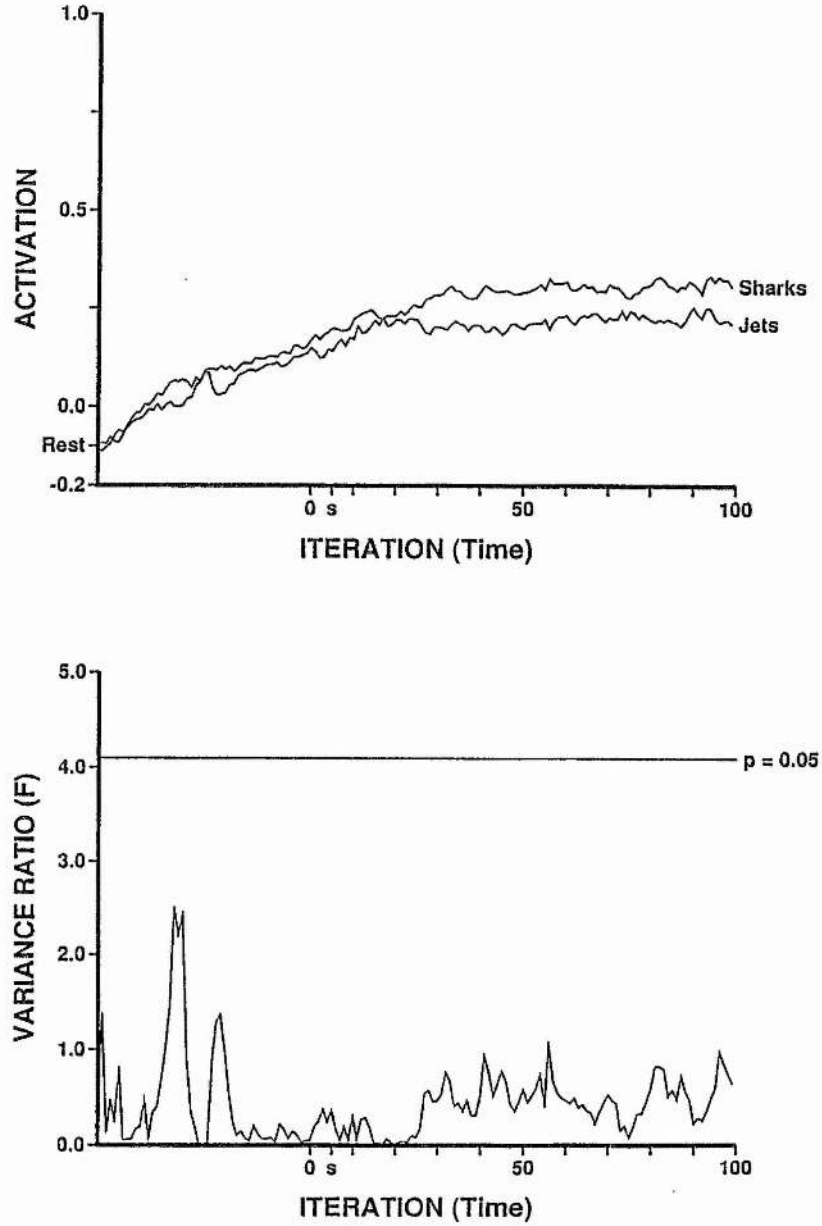


FIGURE 4.7. AVERAGE RESPONSE AND DISCRIMINATION WITH 5% GAUSSIAN NOISE. Same as figure 4.5 except that 5% Gaussian noise was used in the simulations.

individual simulations. Furthermore the discrimination was not reliable until after the activation level of both nodes was substantially above the resting level of the nodes in the noisy network.

It is interesting to note that when 5% Gaussian noise was introduced the network would perform qualitatively differently. One or other node frequently reached and maintained high activation levels before the onset of the input "stimulus": because of this the 'Jets' node would sometimes "win" against the 'Sharks' node (see figure 4.6). Under these conditions, the average over 20 repeated runs looks very different from the noise free simulation and indeed this is reflected in the discrimination analysis which indicates that a significant difference in activity levels was not reached (figure 4.7).

Although the system becomes inherently unstable (when run with no input, one or other node quickly becomes the "winner") the steady state reached by the network is itself stable. That is to say that once one of the two monitored nodes reaches a high level, then it remains at a high level and suppresses the other competing node (for at least 500 computational cycles, figure 4.8).

The use of Gaussian noise could of course introduce occasional large fluctuations in individual nodes, and the effects seen (in particular when using 5% Gaussian noise) could have been due to these "freak" perturbations. To check whether or not this was occurring when 2% Gaussian noise was used, the simulations were run using noise randomly selected from the range -2 to +2% of the minimum to maximum activation levels (thereby removing "freak" activation level changes). Figure 4.9 shows two examples of individual runs using 2% linear distributed noise. As is evident from comparison between figures 4.9 and 4.4 the linear noise did not produce such large fluctuations during normally stable periods from one cycle to the next. However the slower, larger fluctuations in the activity levels of the nodes were very similar to each other.

The average activation levels of the 'Jets' and 'Sharks' nodes and statistical reliability of discrimination from 20 runs using 2% linear noise are shown in

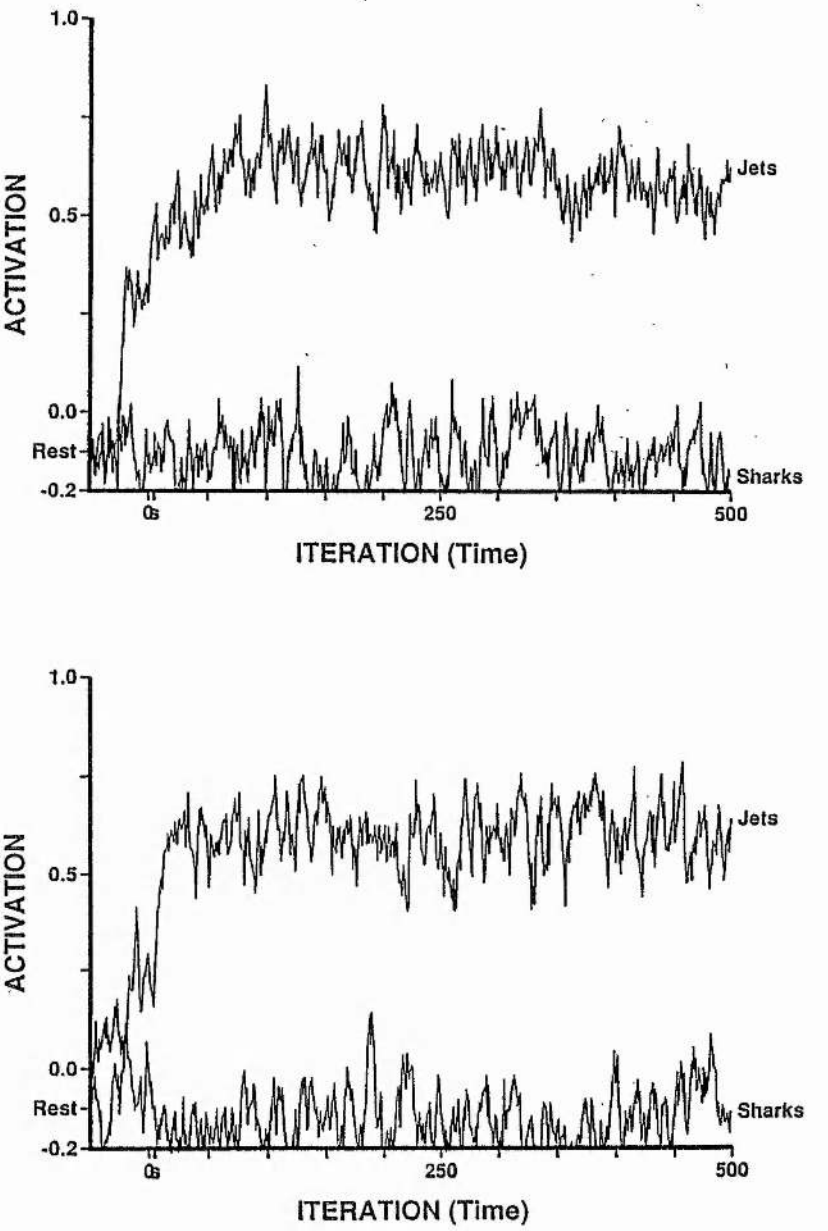


FIGURE 4.8. STABILITY OF THE MODEL WITH 5% GAUSSIAN NOISE. Two examples showing that the model, with 5% Gaussian noise reaches a state that is maintained even though the model is unstable at its nominal resting state. In both examples the 'Jets' node becomes higher than the 'Sharks' node, although this was not always the case (see figure 4.9 upper)

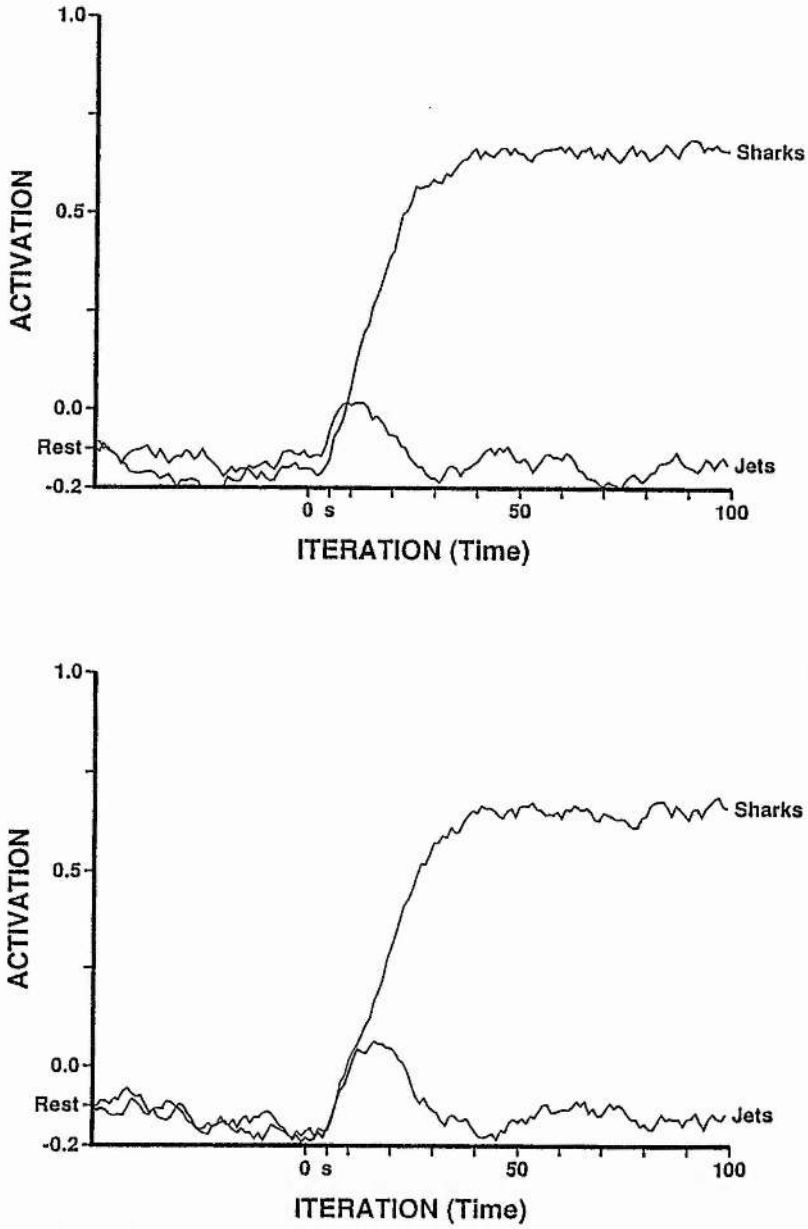


FIGURE 4.9. RESPONSE OF THE MODEL WITH 2% LINEAR NOISE. Two examples of the activation levels of the 'Jets' and 'Sharks' nodes of the network when there is 2% flat distributed noise added internally to the model. The activation level of the 'College' node is set to 1 at iteration 0. The tick marked 's' indicates the iteration where in the noiseless simulation the activation of the 'Jets' and 'Sharks' nodes first separate.

figure 4.10. The results are similar to those obtained using Gaussian noise in that the discrimination between the activity of the 'Jets' and 'Sharks' nodes arise slightly later than it appears in the noise free model. This therefore suggest that the effects were not due to "freak" changes in node activation levels. Note that again discrimination arises after the average activity has risen from the resting activation level. In the example shown in figure 4.10, discrimination becomes statistically reliable only 2 cycles after the time of activation level change in the noise free simulation. Reliable discrimination was not always evident so quickly after the time of activation level change in the noise free simulation, although it is probable that the discrimination is delayed more when using Gaussian noise (see later).

DISCUSSION

The simulations outlined above indicate that the model used shows sensitivity to internal noise (as opposed to noise added to the inputs). This is to say that when noise is added to the processing elements within the model (i.e. the nodes), then the behaviour of the model to a given input changes compared to the noise free situation. The level of noise within the network can cause two qualitatively different effects.

At lower levels of noise (2% Gaussian), the network behaves very similarly to the noise free situation except for the small fluctuations around the expected values. With 2% linearly distributed noise addition, the network reliably reaches towards the same equilibrium state that is achieved without noise (for several hundred simulations). The story is slightly different with 2% Gaussian noise, since the network does occasionally (approximately 1 in 50 runs) become stable with the 'Jets' node high and the 'Sharks' node low. However the rarity of this occurring, and the fact that results using linear noise were consistent with the

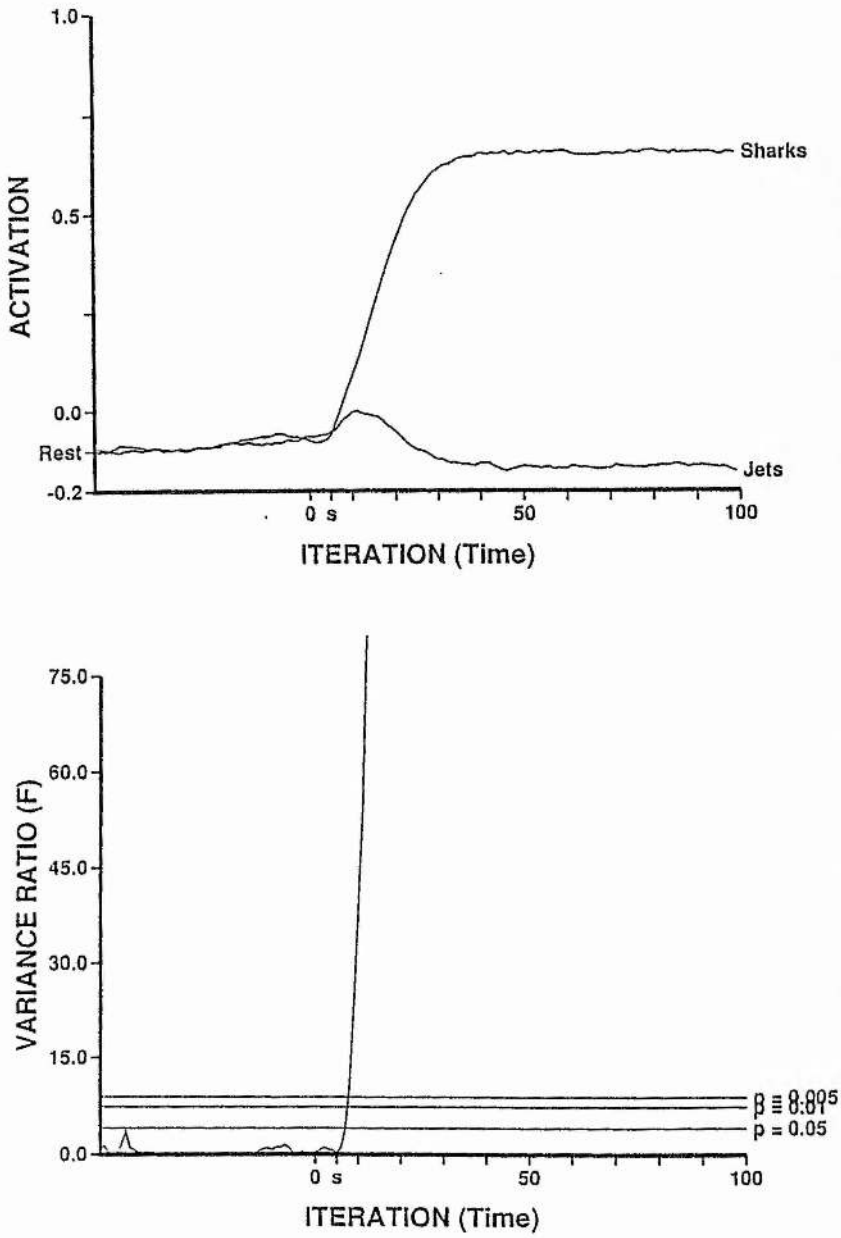


FIGURE 4.10. AVERAGE RESPONSE AND DISCRIMINATION WITH 2% LINEAR NOISE. UPPER: AVERAGE 'JETS' AND 'SHARKS' ACTIVATION LEVELS. The average activation levels of the 'Jets' and 'Sharks' nodes of 20 repeated runs when there is 2% linear distributed noise added internally to the model. The activation level of the 'College' node is set to 1 at iteration 0. The tick marked 's' is the iteration where in the noiseless simulation the activation of the 'Jets' and 'Sharks' nodes first separate. **LOWER:** The discrimination measure (F value) between the activation levels of the 'Jets' and 'Sharks' nodes in the upper portion of the figure. The F values reaches values in excess of 20,000 by the twentyth iteration (not shown). The horizontal lines indicate probability values of 0.05, 0.01 and 0.005. The tick marked 's' indicates the iteration where in the noiseless simulation the activation of the 'Jets' and 'Sharks' nodes first separate.

2% Gaussian noise simulations, suggests that in both cases the results reflect the properties of the network.

For Gaussian distributed noise levels of 5% (and above) the network no long has an equilibrium state with the activation level of nodes around their resting level. Instead the network reaches an equilibrium state with one of either the 'Jets' or the 'Sharks' node high, even if the input to the network is such that, under noise free conditions, the 'Sharks' node is always the node to reach high activation levels.

It is possible that the introduction of small amounts of noise (2%) produces slight changes to some of the computational aspects of the network (i.e. the network's input-output mapping has changed) that are not as clear as the massive change with greater noise levels (5%). While this has not been extensively tested there is reason to think that when 2% noise is added (either linear or Gaussian distributed) this is not the case. The Gaussian noise level added to the system was from the normal distribution having 0 mean and 2% standard deviation. If the noise only adjusted the activation levels of each node independently of its computational function then significance between activity levels should be reached approximately when the difference between the activity levels was approximately 5%. [The expected difference is found from the standard deviate scores, using p=0.05 critical value (1.96) and the standard deviation of the distribution of the noise, here 2% of the difference between maximum (1.0) and minimum (-0.2) possible activation levels, giving 1.96 * 2% * (1 - (-0.2)) = 4.7%.] This was found to be the case for the simulations using 2% noise (both under linear and Gaussian distributions).

The main conclusion from these simulations is that, as predicted, discrimination between the nodes is slow to emerge. In both the 2% linear noise and the 2% Gaussian noise situations the average activation level of both nodes has clearly started to rise and is substantially above the "pre-stimulus" resting

levels before there is a statistically reliable difference between the activation levels of the nodes.

It is argued below that this slow rise phase and accompanying slow discrimination is due to the architecture of the network and not a peculiarity of the particular network that was used. Of course, the slow emergence of discrimination would not be found in a noise free (deterministic) network.

The effects of lateral inhibition in the network during the simulation in the noise free network can be seen shortly after the activation level of the 'Sharks' node exceeds threshold. Once threshold has been reached, the 'Sharks' node starts to exert an inhibitory effect on the 'Jets' node. The 'Jets' node will therefore always receive more inhibition than that which the 'Jets' node itself sends to the 'Sharks' node. This is the main cause of the observation that the 'Sharks' node, once it has an activation level higher than the 'Jets' node *always* maintains a substantially higher activation (in the noise free system).

The "contrast enhancement" effect is increased by similar processes happening in the other pools due to the (direct and indirect) feedback excitation the 'Sharks' node supplies to the other pools. For instance, within the association center, the nodes linking properties of the members of the Sharks gang will be inhibiting those nodes associated with members of the Jets gang. As the difference in activation levels of the 'Sharks' and 'Jets' nodes increases the supportive, excitatory input to the 'Jets' node will decrease. This in turn means that they will be feeding proportionately greater excitation to the 'Sharks' node. The differential activation levels of the association nodes will themselves feed into the other pools, propagating the effects.

The simulations reported here are supportive of the notion that the dynamics of PDP network discrimination are a consequence of the network architecture and will therefore be present in any network that performs "contrast enhancement" computations on its inputs using either feedback or lateral connections.

In summary, the PDP model simulations described above predict that systems using feedback and lateral connections to "contrast enhance" differential activation levels should show a slow initial rise phase, and gradual emergence of discrimination. When noise was added the average activation level of both nodes clearly started to rise and was substantially above the 'pre-stimulus' resting levels before there was a statistically reliable difference between the activation levels of the nodes. PDP networks are often used as models of brain function. Although the class of model used above (interactive activation competition) is normally used for models of high level cognitive processes (e.g. word recognition, McClelland and Rummelhart 1986, and face recognition, Burton et al. 1990) the predictions are based on architectural considerations and will therefore apply to any PDP model with these types of connections (e.g. feedback and lateral).

Comparison with the macaque visual system

It was argued in the previous chapter that the clear discrimination that is present in the first 5 ms of the response of the macaque visual system cannot rely on lateral or feed-back processing occurring either within STPa or during the earlier cortical processing areas because such processing would delay the response latencies beyond 70 ms.

It is worth restating that the behavioural 'decisions' about visual attributes in the image can be based on activities at any of these processing stages: the particular stage used should be task dependent. In other words any behavioural decision should be made as soon as the relevant information for that decision is available. This means that processing does not have to pass through the whole system before behavioural decisions and actions can be taken. Such cascade processing would depend on outputs from each area to sub-cortical systems, rather than proceeding with the hierarchical processing outlined above.

One of the aims of the simulations was to generate predictions of the dynamics of PDP models to allow comparison with data obtained from

physiological recordings. Throughout, an implicit assumption has been made about the nature of the nodes in the model: that they can in some way be compared to single brain cells. To this effect, the discussion has assumed that the activation level of the nodes is directly comparable with observing activation levels of neurons, either singly, or as a population. In this assumption, the activation levels of the nodes prior to input initiation has been related to the "pre-stimulus" period of physiological recordings. In terms of the simulated model however information is only transferred to other nodes once the activation level has exceeded zero. At first glance this might seem an untenable position, yet it must be the case that a similar situation exists within the brain. Recordings from all brain regions show that neurons have spontaneous activity. This means that every now and again neurons give a response (typically an action potential). If there were not some thresholding mechanism, then these random signals would be transmitted to the next stage within the brain and the owner of these cells would be constantly sensing at random intervals stimuli, emotions etc. As subjectively we know this not to be the case, then it is reasonable to regard the non-communicating resting level of the nodes in a model as comparable to the resting level (i.e. spontaneous activity) of neurons, and that detection of a change in the activity of nodes is comparable to the increase in activity of a neural response. Note that this also allows for a high level threshholding mechanism operating in the brain. Indeed Gochin et al. (1991) found that up to 20 or 30 simultaneously active inputs were needed before cells in infero-temporal cortex would spike (a temporal as well as magnitude based thresholding system). The appropriate comparison is therefore between the observable activation levels of nodes and neurons, not at what activation level is the information passed to the next processing element.

Comparison of figures 3.6 and 4.5/4.10 shows that the fluctuations in the pre-stimulus period for the averaged response (Figure 3.6) are much greater than those seen in the model simulation (Figures 4.5 and 4.10). This implies that the

faster discrimination onset observed in the biological system is not due to a lower noise level than was present in the model. These observations confirm the predictions made in the previous chapter and again suggest that the macaque ventral pathway processes visual information in a predominantly feed-forward manner.

CHAPTER 5

MOTION PROCESSING IN STPa DIRECTIONAL TUNING

(Oram et al. 1993, *Exp. Brain Res.* 97:274-294)

INTRODUCTION

Visual information processing in the cerebral cortex of primates appears to have at least two major divisions, one analysing motion the other analysing form. One stream runs ventrally from occipital cortex into the temporal lobe and is thought to be involved in the analysis of visual pattern and recognizing the form of objects (Ungerleider and Mishkin 1982). A second stream projects dorsally into the parietal cortex. This pathway has been postulated to be concerned with the spatial position of objects (Ungerleider and Mishkin 1982) and visuo-motor coordination (Goodale and Milner 1992). Since this dorsal pathway involves areas in the posterior part of the superior temporal sulcus which contain cells almost invariably sensitive to motion, this stream of processing has also been dubbed the motion pathway (De Yoe and Van Essen 1988).

The upper bank of anterior sections of the superior temporal sulcus (STS) contains an area, STPa (anterior superior temporal polysensory region, also known as areas TPO and PGa after Seltzer and Pandya 1978), which is a high level visual processing area receiving input from both ventral (form) and dorsal (motion) processing streams (Felleman and Van Essen 1991; Young 1992). The majority of studies in this region have concerned the selectivity of cells to static pattern information and the analysis of complex biologically significant objects [e.g. the form of hands (Gross et al. 1972), faces and bodies (Bruce et al. 1981;

Desimone et al. 1984; Baylis et al. 1985; Perrett et al. 1982, 1984, 1992; Young and Yamane 1992; see also chapters 1 and 3). Other studies in STPa have described cells selective for complex body movements including hand actions, patterns of walking and head and limb articulation (Bruce et al. 1981; Perrett et al. 1985b, 1989, 1990a, 1990b; Mistlin and Perrett 1990; Hasselmo et al. 1989). These latter cells are interesting because they may indicate the integration of the form and motion streams of information at the cellular level (see chapters 7-9).

Despite the preponderance of cells in STP with complex selectivity there are also large numbers of cells sensitive to motion but showing no apparent sensitivity to form in the same area (Bruce et al. 1981; Perrett et al. 1985b; Hikosaka et al. 1988; Mistlin and Perrett 1990). The functional significance of this cell population is unclear.

Bruce et al. (1981) distinguished three main types of direction selective cells in STPa; those sensitive to movement in the fronto-parallel plane, movement in depth and radially symmetric movement about the centre of gaze. These types of cell responses are very similar to those found in the medial superior temporal area (MST) in the posterior section of the STS (Duffy and Wurtz 1991a; Tanaka and Saito 1989; Tanaka et. al. 1989). Other STP cells exhibit less directional selectivity responding to multiple or all directions of motion, a property not reported for earlier visual motion areas. A variety of less common STPa cell types have also been reported which were sensitive to rotation, jerky motion, or appearance or disappearance from the visual field (Bruce et al. 1981; Perrett et al. 1985b). Bruce et al. (1981) reported that the majority of the STPa neurons displayed little or no form specificity. Perrett et al. (1986b) also found one quarter (84/335) of the STP motion sensitive cells lacked form selectivity.

One possible function of these motion sensitive STPa cells lacking form selectivity might be to contribute to the properties of cells conjointly sensitive to form and motion. Intuitively, these could be created by combining the outputs of cells sensitive to the static form with the outputs of cells sensitive to direction of

motion. In order to evaluate this scheme more information is needed about the motion sensitive STPa cells lacking form selectivity. Such information could also clarify their relation to motion processing in regions within the dorsal pathway.

Motion pathways

In order to understand the motion sensitive properties of cells in STPa it is useful to review motion processing in earlier cortical regions of the motion pathway. In the macaque monkey the cortical processing of motion information involves a hierarchical series of steps through magnocellular layers of the lateral geniculate nucleus, the upper and lower portions of layer 4Cα of V1, the thick stripes of area V2, the middle temporal area (MT or V5), area MST, an area on the floor of the superior temporal sulcus (FST) and posterior sections of the STP (STPp).

Whilst areas V1 and V2 contain neurons selective for both static and moving visual stimuli in roughly equal proportions, areas in the posterior sections of the superior temporal sulcus (MT, MST, FST and STPp) contain a very high proportion of motion selective cells (Albright et al. 1984; Duffy and Wurtz 1991a; Hikosaka et al. 1988; Maunsell and Van Essen 1983; Movshon et al. 1985; Rodman and Albright 1989; Zeki 1974a).

In V1, V2 and MT there is clear retinotopic mapping but, in MST, FST and STPp this mapping does not appear to be present (Gattass and Gross 1981). Further, in MT and MST, receptive fields show little or no relationship between eccentricity and receptive field size (Komatsu and Wurtz 1988a; Tanaka et al. 1986). Like V1 and V2, both MT and MST do show evidence of a columnar organization with similar cell displaying similar motion sensitivity occurring in close proximity (Albright et al. 1984; Saito et al. 1989). Proximity of similar cell types has also been reported in STPa (Harries and Perrett 1991; Perrett et al. 1984, 1985b).

Receptive field size increases going up the motion processing stream. For instance, cells in MT have receptive fields approximately five times the size those in V1 (Mikami et al. 1986b). In the next cortical area of the motion processing hierarchy, MST, the receptive field size of increases again, typically extending into the ipsilateral hemispace (mean square root size range 45--53 degrees, Tanaka and Saito 1989; 62--64 degrees, Duffy and Wurtz 1991a). From MST to STPa there is again an increase, with receptive field size of some 80% of the neurons in STPa covering nearly all of the visual field (median size 150 degrees horizontal, 105 vertical, Bruce et al. 1981).

Cells of V1 and V2 with directional selectivity show a preferred direction of motion that is perpendicular to the orientation of bar stimuli. In area MT some 30% of cells show selectivity for a direction of motion parallel to the preferred bar orientation, allowing the area to code the global direction of an object's motion independent of its local contour orientations (Albright 1984; Snowden et al. 1991). MST neurons, particularly those in the dorsal region, MSTd, prefer motion over a wide field (Duffy and Wurtz 1991a, b; Komatsu and Wurtz 1988a,b; Tanaka & Saito 1989, Tanaka et al. 1986, 1989). Further, most neurons in MST show no response to self-induced retinal motion resulting from eye movement, a property not observed as frequently in neurons of V4 and MT (Duffy and Wurtz 1991b; Erickson and Thier 1991; cf Galletti et al. 1990), are relatively insensitive to speed or dot density of moving patterns (Duffy and Wurtz 1991b), but are sensitive to disparity (Roy and Wurtz 1990; Roy et al. 1992).

Tanaka and Saito (1989) have suggested that the sensitivity to wide field motion in MST has a role in maintaining visual stability during self-motion and hence to control of posture. Despite the size of their receptive fields, STP cells do not require large field stimuli. Thus as the hierarchy of motion processing areas is ascended towards the STP, receptive fields increase in size and cells become less responsive to self-induced motion (Hietanen and Perrett 1992). The functional

role of motion processing within the STP, however, remains unclear particularly for those cells lacking form selectivity.

Form processing

The manner in which static form is processed within the STPa has been more fully characterized and may provide insight into motion processing in the same area. The vast majority of STPa cells sensitive to the static form of objects display selectivity for perspective view (Desimone et al. 1984; Perrett et al. 1991, 1992). The distribution of view tuning displayed an interesting inhomogeneity; namely in the horizontal plane of rotation, statistically more STPa cells were found with optimal views of the head and body close to the front and side views of the head than to intermediate views (Perrett et al. 1991, 1992). This uneven distribution of optimal views amongst STPa neurons supports theoretical models of recognition whereby a small number of 'characteristic' views of the object are selectively represented in the nervous system (Koenderink and van Doorn 1979).

One of the main aims of the present study was to examine the directional selectivity of STPa cells and to determine whether the distribution of directions preferred by cells was continuous, or whether particular directions were preferentially represented. Since other classes of STP cell are selective to head or body view *and* direction of motion, the preferential analysis of particular directions might facilitate the integration of the two types of information. Studies of V1 indicate a slight bias for coding horizontal and vertical directions of motion (Mansfeld and Ronner 1978; DeValois et al. 1982). Studies of MT and MST have not noted any strong bias in the distribution of optimal directions of motion (e.g. Albright 1984). Despite this apparent absence of preferential tuning to particular directions there is some evidence for its appearance in STPa. In a preliminary report optimal direction of STPa motion sensitive cells appeared to coincide with cartesian axes (up/down, left/right and towards/away, Perrett et al. 1985b, 1990a,b) but no systematic study has yet been made of direction tuning in STPa.

Thus there is a need for quantitative study of tuning to determine whether processing of direction of movement is similar to the processing of static view.

A second aim of the study was to define the breadth of tuning for direction. This is likely to be related to the distribution of preferred directions. With broad tuning (45 degrees, 1/2 width at 1/2 height) four populations of cells tuned to directions 90 degrees apart could represent all directions of motion in a plane, with narrower tuning of, say 22.5 degrees, 8 populations would be needed. Of course, finding broad tuning for direction in the STP does not itself guarantee an uneven distribution of directions preferred by cells. For example cells in MT are found tuned to a continuous range of directions yet tuning appears relatively broad (Albright 1984). Knowledge of the breadth of tuning is also important for comparison with motion processing in other areas.

The time course of neuronal responses to movement was also examined in the present study since this information can help specify the likely source of visual input to the motion sensitive STP cells.

METHODS

Two female (wt 4 Kg) and three male (wt 5--10 Kg) rhesus macaque monkeys were used. The monkeys are referred to as F, J, B, D and H.

Testing procedure

The standard behavioural task and recording techniques were used. Each cell recorded was first subjected to exploratory testing involving the presentation of a variety of static and moving objects. Testing associated with other experiments involved presenting tactile, auditory stimuli and up to 8 views of static and walking human bodies. Any hand held stimulus motion was started before the shutter opened and continued until after the shutter closed. Video disk

images of moving stimuli were under the control of the computer and were therefore exactly repeatable. Speed sensitivity was assessed with hand held stimuli using three broad categories: fast (> 30 degrees/sec), medium (10--30 degrees/sec) and slow (< 10 degrees/sec). This ensured that directional tuning was assessed for each cell at an effective stimulus speed. The accuracy of live 3D presentation was assessed by analysis of video recordings of movements of a typical testing protocol. The analysis took the form of marking the object in each frame of the video sequence and storing the X-Y co-ordinates using a Pluto Graphics system (IO Research). Position and velocity profiles could the be calculated for each of the directions tested. It was found that the variation of the mean speed of motion between directions was within +/-- 15%, and the overall range of speeds was within +/-- 30%. Directional accuracy of hand held stimuli was better than +/-- 10 degrees.

A cell which exhibited consistent responses only to moving stimuli or gave preferential responses to stimulus motion was tested for possible selectivity for direction of movement. The cells were routinely tested for six different directions of movement along 3 orthogonal axes (towards, away, up, down, right, left). This testing included moving 3D stimuli in front of the monkey in the preferred direction(s) under strong diffuse room lighting (>800 watt total). The stimuli included human faces and bodies and various hand-held laboratory objects of different shape, size (subtending 1 to > 20 degrees), colour and texture (fruit, tools, boxes, curtains, fur, bodies, etc.). Cells were also tested with moving 2-D stimuli from a video disk library which included simple geometrical images (e.g bars, spots, gratings) as well as complex images of moving bodies. These video stimuli moved in 8 directions in a given plane and allowed precise repetition of stimulus trajectory etc.

If the cell was observed to respond equally to all stimuli tested in the preferred direction(s) it was classified as a non-form, motion sensitive cell (e.g. see Figure 5.1). In some cases the size of the object was found to have an effect

on the responses but, as no other selectivity for features could be established, these cells were also classified as non-form selective. Cells which were found selective for both motion and stimulus form will be reported elsewhere. Speed of stimulus motion was also regularly tested (5 degrees/s to 100 degrees/s) and when found to have an effect the preferred speed was used for all subsequent testing. The effect of the position of motion within the visual space was also examined and again if there was found to be an effect of the position of stimulus presentation the most effective position was used for subsequent testing.

Cells lacking form sensitivity but which showed a tendency to discriminate between moving and static objects were tested with 5 trials of four or eight directions of movement, presented in a computer-controlled and randomized order. Testing was performed in one mode using either real 3D, projected 2D slides or video disk stimuli. Computer-controlled testing protocols enabled data to be subjected to ANOVA and regression analysis on-line.

Data Analysis

Assessment of response magnitude

Cell responses to 4 or 8 directions, static controls and spontaneous activity (S.A.) were compared on line using 1-way ANOVA and post-hoc tests (protected least significant difference, PLSD, Snedecor and Cochran 1980). For cells tested with eight directions multiple linear regression analysis was used to estimate the best relationship between response and 2nd order cardioid function of direction. In effect this calculates the values of the coefficients β_{1--5} of the equation below which produce the highest correlation between response and the angle of motion.

$$R = \beta_1 + \beta_2 \cos\theta + \beta_3 \sin\theta + \beta_4 \cos 2\theta + \beta_5 \sin 2\theta,$$

where R is the response, θ is the directional angle and β_{1--5} are coefficients.

This equation was chosen because it makes very few assumptions about the nature of direction tuning. It also provides a good estimate of tuning of cells with a single preferred direction and cells with two preferred directions approximately 180^{o} apart (e.g. movement left and right). See Perrett et al. (1991 and Appendix) for a detailed discussion.

Where the regression analysis produced a significant ($p < 0.05$) relation between predicted and observed values, the regression equation was used to define: (a) the optimal direction (θmax), (b) the maximum response at this direction (Rmax), (c) the sharpness of tuning (average angle of rotation required to reduce the response to half Rmax) and (d) the angle and magnitude of any second peak in the direction tuning.

Assessment of response time course

The criteria for the onset of cell responses was set at the upper limit of the 95% confidence interval of the pre-stimulus period. The latency was assessed for the responses to the most effective direction(s) and taken as the first of three consecutive (5.0 or 5.2 ms) bins exceeding the onset criteria (Oram and Perrett 1992; chapter 3). For these same cells, the responses in each time bin were normalized so the spontaneous activity was set to 0 and peak response set to 1.0. Averaging responses across cells gave the population response profile.

Once the latency estimate had been made, firing rates were calculated for each cell for the first, second and fifth 100 ms periods after response onset. The firing rate during the final 100 ms of the data collection period was also calculated for each cell (800--900 ms after response onset). The peak firing rate was taken as the maximum value of a running 3 bin average of firing rate. The rise time was calculated as the time from response onset to the peak firing rate. The half fall time was calculated as the time from peak to the time when the running 3 bin average fell below (Peak --S.A.)/2. The decay time was calculated as the time from peak to the time when the running 3 bin average fell below the threshold

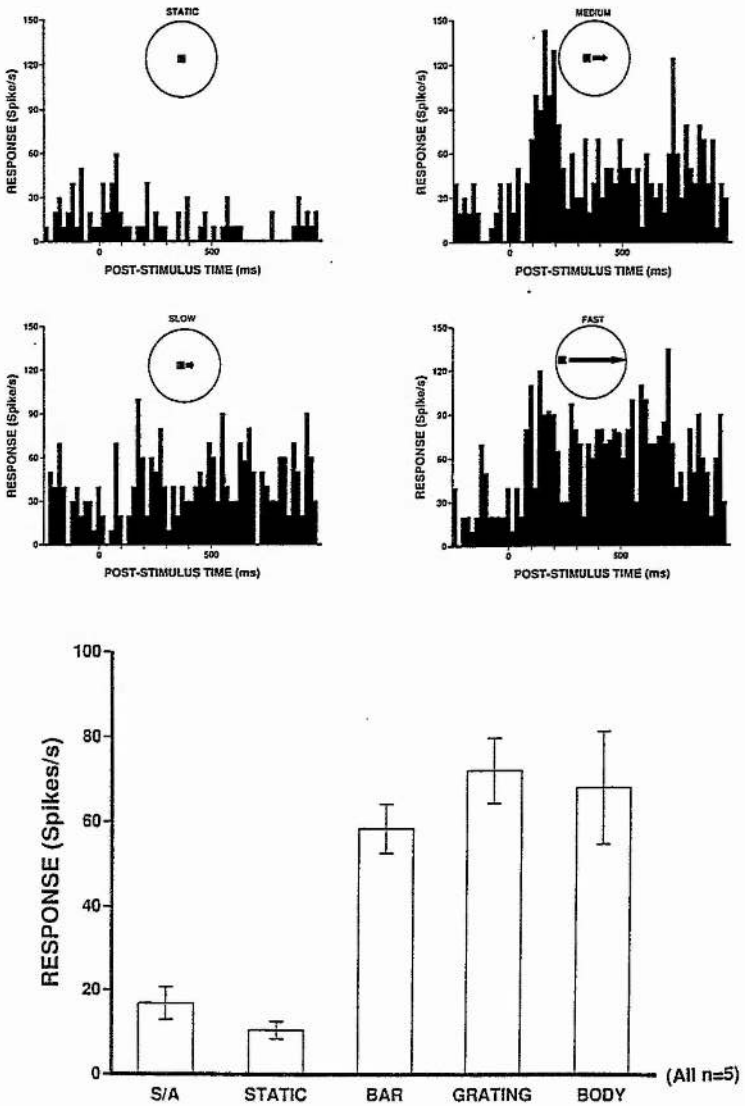


FIGURE 5.1. CELL SELECTIVE FOR MOTION BUT NOT FORM. UPPER: Peri-stimulus time histograms (PSTH) of responses of one cell to static and moving stimuli (5 trials per PSTH, bin width = 20 ms, stimulus: experimenter moving under diffuse strong room lighting. Movement speeds: static, slow, medium, fast approximately 0, 7, 13, 40 degrees/s or 0, 0.5, 1.0, 3.0 m/s at the viewing distance of 4.0 m). Above each PSTH is a schematic depiction of the motion. The circle depicts the visual field through the shutter (the fixation LED is central). The square gives the approximate starting position of the (already moving) stimulus at time 0 when the shutter opened, and the arrow indicates the direction of movement. The tip of the arrow indicates the approximate end position of the stimulus when the shutter closed. The response to any motion produced a response above static stimuli ($p < 0.01$ each comparison). There was also an effect of speed (fast and medium > slow, $p < 0.005$ each comparison). Overall effect of conditions $F_{[4,20]} = 16.6$, $p < 0.0005$. **Lower:** Histogram of response magnitudes of the same cell to 3 different stimuli (bar, subtending 2 degrees, grating, subtending 5 degrees and body, subtending 20 degrees) moving at 20 degrees/s, spontaneous activity (S/A) and static objects (static: bar, grating and body). All three objects produced equivalent responses when moving ($p > 0.2$ each comparison) which were greater than spontaneous activity or the static presentation of the same objects ($p < 0.002$ each comparison). Overall effect of conditions $F_{[4,20]} = 14.9$, $p < 0.0005$).

criteria used for the latency estimate. Finally, the duration was taken as the time from response onset to the first time when the running 3 bin average fell below threshold level. Note that for some cells this was before peak firing rate had been reached.

RESULTS

Cell classification

553 of the visually responsive cells were classified as lacking selectivity for stimulus form but having sensitivity for motion. It should be pointed out that few of these cells showed selectivity for speed or position in the visual space of the motion. Where such selectivity was noted, optimal conditions were used for further testing. Figure 5.1 shows an example of a cell which was selective for the speed of motion (upper) but not the form of the stimulus (lower). Shown above each of the peri-stimulus time histograms (PSTHs) is a schematic representation of the stimulus motion within the area revealed from the open shutter. The LED was situated in the centre of this area. The stimulus was moved by hand at speeds corresponding to 0, 7, 13, 40 degrees/s (+/-- 15%). As can be seen for all motions there was at least a slight response, and above 10 degrees/s the cell produced a clear response. In contrast for static stimuli there was slight inhibition. The lower section shows the mean response (spikes/s) of three objects (a simple bar, grating and a body) moving at approximately 20 degrees/s. As can be seen, the response magnitudes are all equivalent ($p > 0.2$), but substantially greater than the spontaneous activity or the presentation of static stimuli ($p < 0.002$ each comparison).

Cells were screened to check that the response differences to different stimuli were not due to differences in eye position or movements. No relation was observed between responses and eye movements for any of the 43 cells where eye

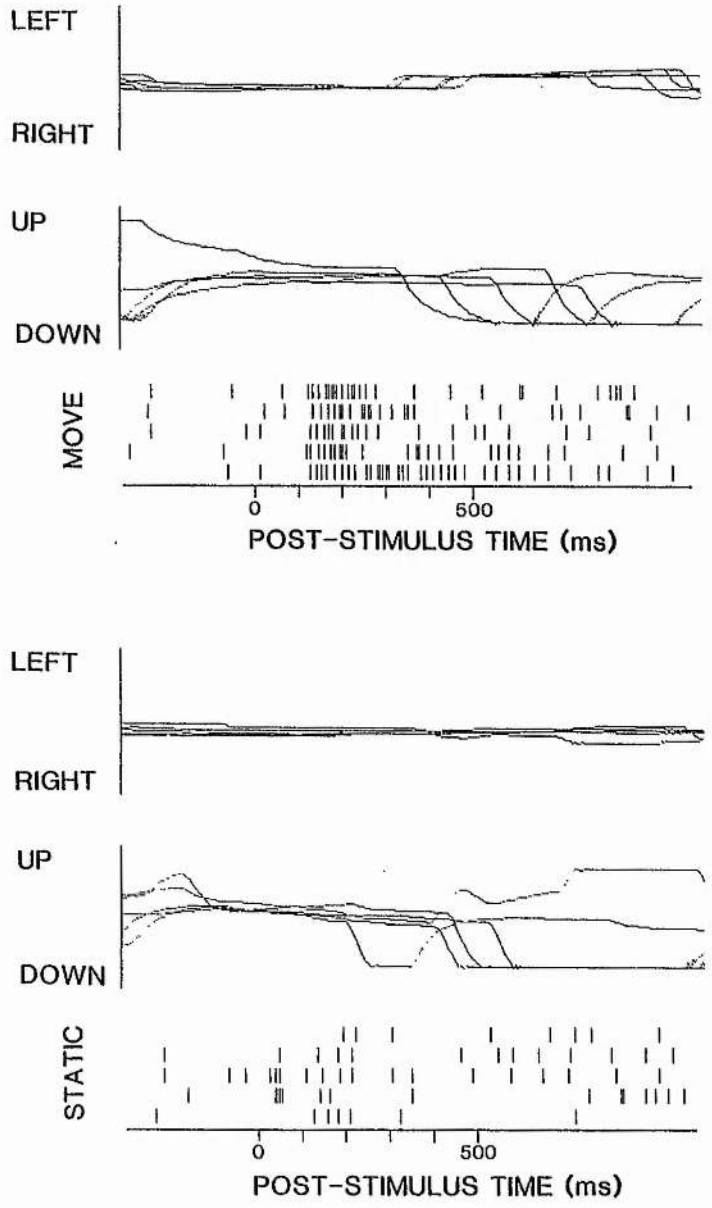


FIGURE 5.2. EYE MOVEMENTS DURING EFFECTIVE AND NON-EFFECTIVE STIMULUS PRESENTATION. UPPER: Eye position (upper) and responses (lower) of one cell to 5 presentations of an effective stimulus. For this cell, movement to the subject's left in the lower half of the visual field elicited a good response (stimulus: hand moved red square under bright diffuse room lighting subtending 3 degrees, moving 10 degrees/s). At the onset of trials the subject fixated a centrally positioned LED to determine its colour. Later in the trials the subject made saccades down to the stimulus. **LOWER:** Eye position (upper) and cell responses (lower) to 5 trials of the same object presented stationary in the lower half of the visual field. The eye movements are comparable for cases shown in the upper and lower sections, yet only when the stimulus was moving was there a cell response.

movements were recorded. Figure 5.2 gives an example of eye position recordings during an effective (moving) stimulus and ineffective (static) stimulus. Recordings show the monkey fixating the position of the coloured LED before stimulus onset (0 ms) and maintaining fixation for at least a further 200 ms. With the stimuli moving to the left, the cell responded at a latency of 110--130 ms regardless of the latency and pattern of subsequent eye movement. The pattern of eye movements elicited to static stimuli was similar, yet the neural response was abolished. As with all cells reported in this study, there was little if any response to static stimuli. The cell was also tested with movements to the right and towards the subject (not shown in Figure 5.2). These directions of motion produced different patterns of eye movement after the data collection period but similar neuronal responses. Importantly the monkey was fixating during the sample period on which the analysis was based (100--350ms post-stimulus onset) for all except one trial. This pattern of maintaining fixation during the sample period was observed for almost every trial where eye movements were recorded. On the occasional individual trials when the monkey broke fixation before the 350 ms time there was no clear evidence of a change in the response, either before, during or after the saccade. As the performance of the subjects at the LED colour discrimination task was consistently high with multiple licks, there is no reason to suspect that fixation patterns differed for the other tested cells.

As mentioned above, all the cells were routinely tested for six different directions of movement along 3 orthogonal axes. If a cell was found responding preferentially in only one of these directions it was classified as a *unidirectional* cell. Based on the routine screening testing, Table 5.1 presents the distribution of the preferred directions of unidirectional cells recorded from all five subjects. 216/553 (39%) non-form selective motion sensitive cells were classified as unidirectional. *Bidirectional* cells were classified as cells which showed roughly equal responses to two directions with responses in between which were substantially weaker. 23/553 (4%) non-form selective motion sensitive cells were

classified as bidirectional. Finally, the remaining cells showed approximately equal responses to motion in many or all directions and were classified as pandirectional (314/553 or 57%). This class of cells may have included cells displaying the radial type of motion sensitivity described by Bruce et al. (1981).

Table 5.1

	Preferred Direction						
Subject	U	D	R	L	+	-	Total
B	1	3	10	10	4	8	36
F	26	13	4	0	39	13	95
D	6	7	9	6	15	5	48
H	1	1	0	0	2	0	4
J	7	5	1	7	9	4	33
Total	41	29	24	23	69	30	216

Table 5.1. Distribution of STPa cells tuned to different directions. Classification of preferred directions of cells from qualitative assessments in 5 recording subjects (B, F, D, H and J). Abbreviations: U = up, D = down, L = left, R = right, + = towards and -- = away.

After initial directionality screening, the directionality of 43 cells was tested with 8 directions of motion in a given plane. Three cells were tested twice in the same plane to asses reliability of testing, 8 cells were tested in two different planes and one cell was tested twice in one plane and a third time in a second plane giving a total of 56 regression analyses. Of this total of 56, 50 (89%) were found to give a significant relation between response and the second order

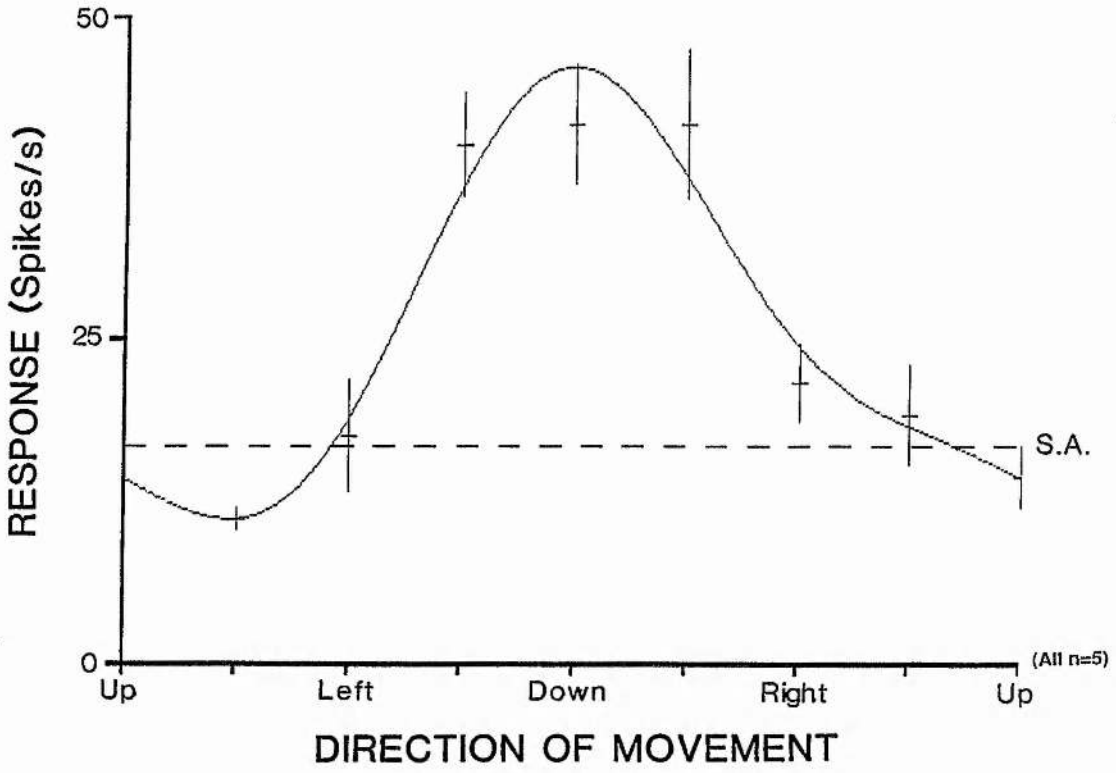


FIGURE 5.3. RESPONSES OF A UNIDIRECTIONAL MOTION SENSITIVE CELL. The mean responses (+/- 1SE) are illustrated for one cell to 8 directions in the fronto-parallel plane (stimulus: hand held light bar in blackout conditions swept across visual field at approximately 50 degrees/s). Direction is expressed as the angle of rotation from upwards (0=up, 90=left, 180=down, 270=right). The curve is the best fit cardioid function, relating response to direction (R^2 = 0.675; $F_{[4,35]}$ = 18.2, $p < 0.0005$). The dashed line denotes spontaneous activity (S.A.). Responses to movement downwards (down and left, down and right and straight down) were significantly greater than movements in other directions and S.A. ($p < 0.005$ each comparison) but were themselves equivalent ($p > 0.75$). Overall effect of conditions $F_{[8,36]} = 10.6$, $p < 0.0005$.

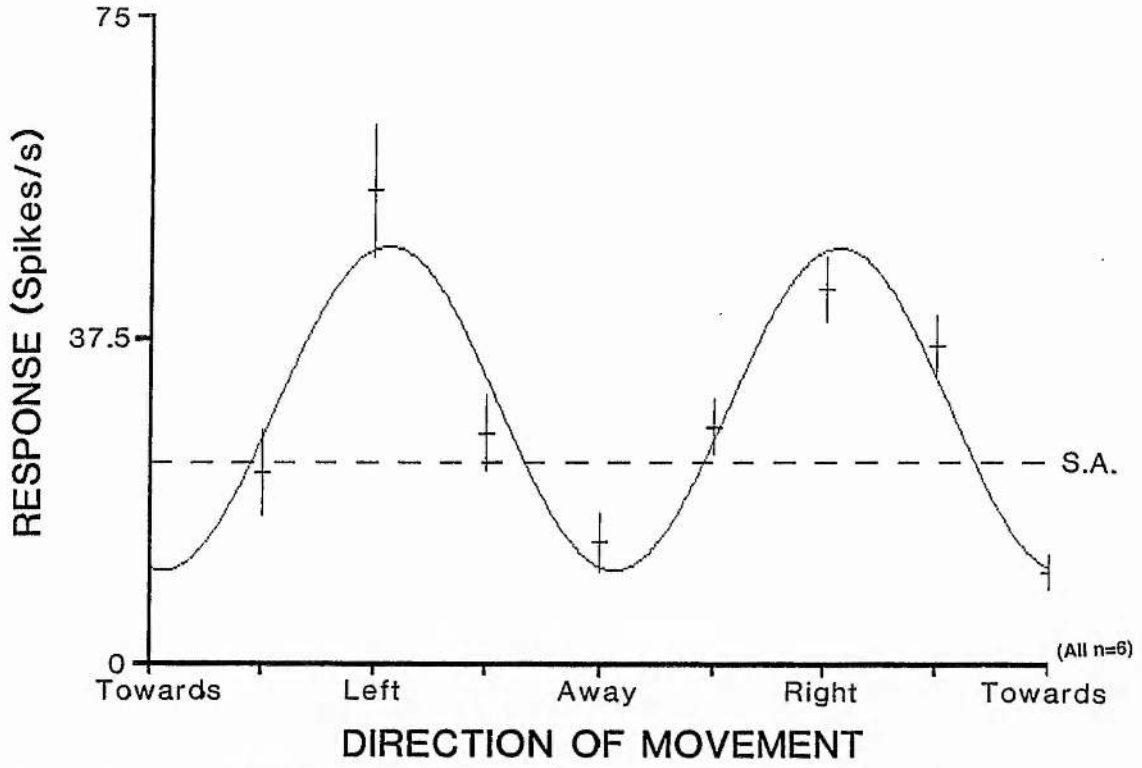


FIGURE 5.4. THE RESPONSES OF A BIDIRECTIONAL MOTION SENSITIVE CELL. Responses (mean +/- 1SE) of one cell to 8 directions in the horizontal plane (stimulus: experimenter walking 3m/s under strong diffuse room lighting, at a mean distance 2m). Direction is expressed as degrees of rotation away from movement towards the subject (0=towards, 90=left, 180=away, 270=right). The curve is the best fit cardioid function, relating response to direction (R^2 = 0.604; $F_{[4,43]}$ = 16.5, $p < 0.0005$). The dashed line represents spontaneous activity (S.A.). The cell responded to movements left or right more strongly than to movements towards or away from the subject ($p < 0.0005$ each comparison). The responses to movements left and right were statistically indistinguishable (p = 0.09). Overall effect of conditions $F_{[8,45]} = 9.1$, $p < 0.0005$.

cardioid function of direction of movement. Of these 50 cells 19, 18 and 9 cells were studied in the horizontal, fronto-parallel and sagittal planes, respectively. [The 3 cells retested in the same plane, 8 cells tested in two planes and the 1 cell which was both retested in the same plane and tested in a second plane all gave significant fits].

The responses of 32 of the direction selective cells followed a *unimodal* pattern (unidirectional cells), with one direction evoking the optimal response. An example of a unimodal or unidirectional cell is given in Figure 5.3. For this cell, movement with a downward directional component elicited a strong response, whereas movement to either side of the subject or upwards produced responses no different from spontaneous activity (S.A.).

Five direction-selective cells were classified as bidirectional because their responses to two directions were both significantly ($p < 0.05$) higher than intervening directions. Figure 5.4 shows responses of a bidirectional cell selective for motion to the subject's left and right. For all bidirectional cells in this study the two preferred directions were approximately 180^{o} apart, even though it could have been possible to find a cell with two preferred directions only 90 degrees apart. The criteria for classification as bidirectional used here were fairly stringent and a further 3 cells, classified as unidirectional, showed a degree of bidirectional direction tuning, in that their response to a second or minor direction was greater than half the response to the optimal direction (with other intervening directions evoking less than half the maximal responses). These 5 bidirectional cells were unlikely to be the radial type reported by Bruce et al. (1981) as the radial motion that occurred when the stimulus was moved in the directions intervening between optimal directions produced only weak responses. Bidirectional cells responding to movements left and right did not respond to movement up or down, when in all cases there was equivalent radial motion.

As already noted, the majority of non-form selective motion sensitive cells recorded were responsive to movement of an object in any direction. Some of

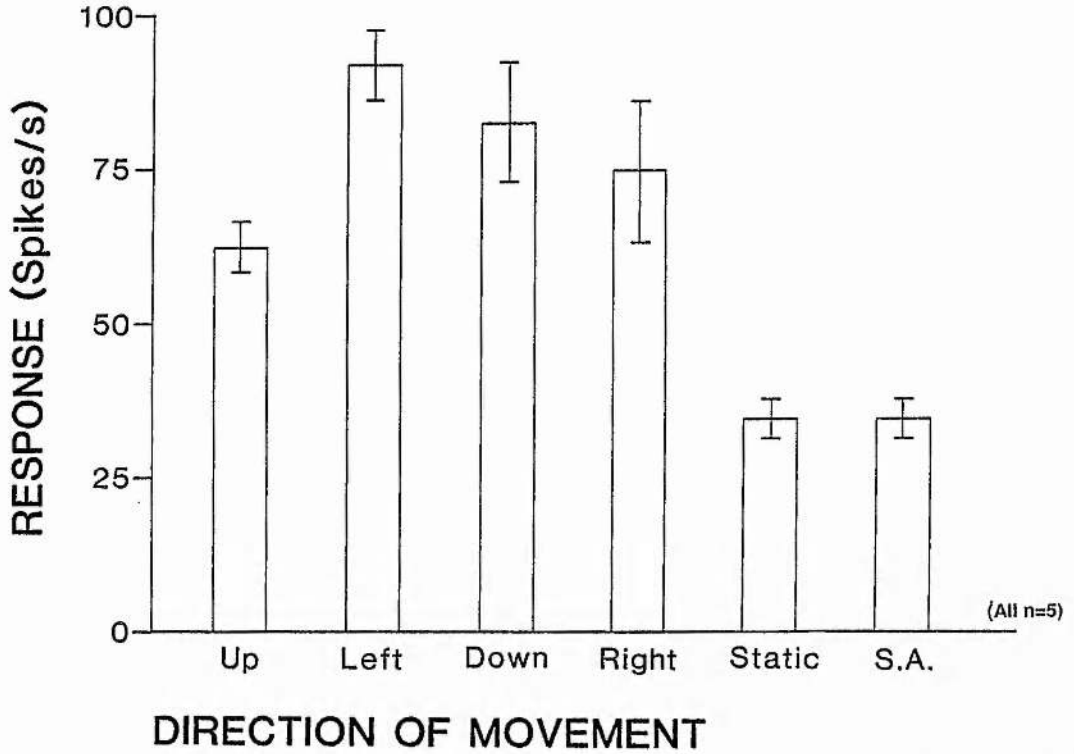


FIGURE 5.5. THE RESPONSES OF A CELL RESPONSIVE TO MULTIPLE DIRECTIONS OF MOTION. Responses of a single cell to 4 directions in the fronto-parallel plane (stimulus: hand held green box subtending 30 X 50 degrees, moving at 40 degrees/s). Movement in any of the tested directions gave a response that was greater than spontaneous activity (S.A.) and static views of the same object ($p < 0.01$ each comparison). The cell responses, however, showed slight selectivity for direction of movement with a significantly greater response to movement to left and down than movement up ($p < 0.05$). Overall effect of conditions $F[5,24] = 12.0$, $p < 0.0005$.

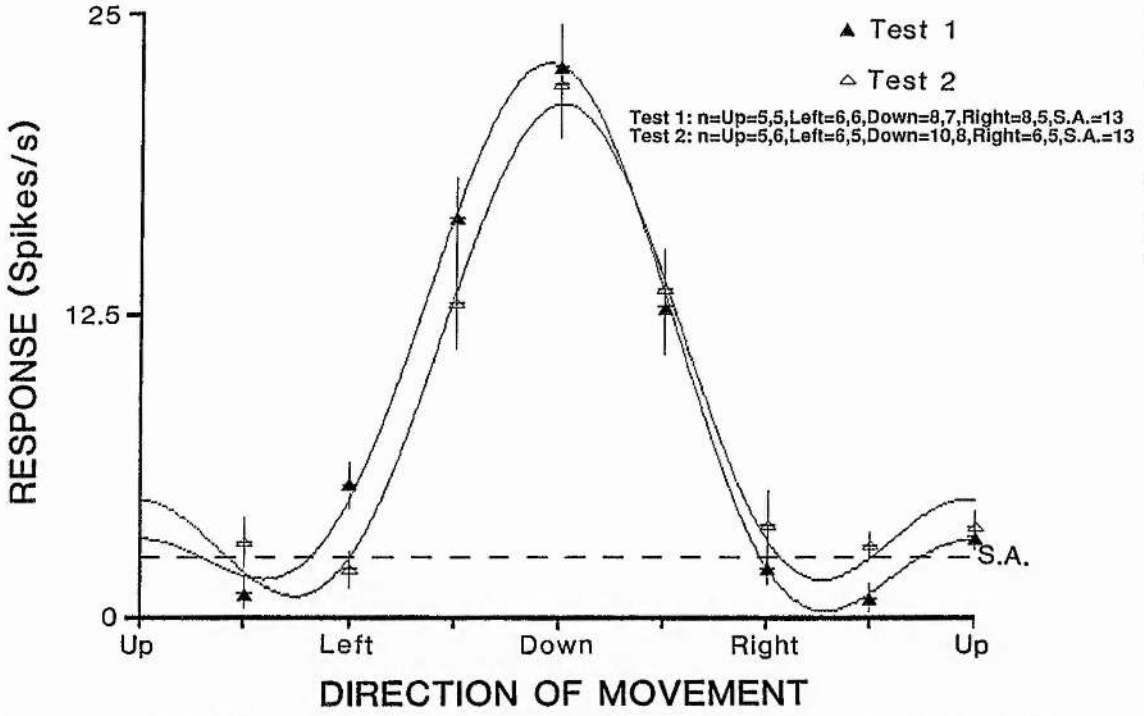


FIGURE 5.6. TEST-RETEST RELIABILITY FROM REPEATED DIRECTION ANALYSIS. Responses and regression curve of one cell to two tests of 8 directions in the fronto-parallel plane (see text for stimulus details). The cell showed strong responses to movement downwards in both tests. Estimated maximal responses are at 174^{o} and 180^{o}. Overall effects, Test 1: ANOVA $F_{[8,54]} = 54.1$, $p < 0.0005$; Regression $R^2 = 0.882$, $F_{[4,45]} = 84.4$, $p < 0.0005$. Test 2: ANOVA $F_{[8,57]} = 25.7$, $p < 0.0005$; Regression $R^2 = 0.751$, $F_{[4,48]} = 36.2$, $p < 0.0005$.

these cells (e.g. Figure 5.5) also exhibited weak directional selectivity. For the cell illustrated in Figure 5.5 movement of an object in any direction in the fronto-parallel plane caused the cell to respond, whereas the same object held stationary elicited no response. The responses to all movements were not equivalent; movement to the subject's left and downwards was significantly greater than movement upwards.

Whilst the majority of the cells were tested only once, 4 cells were subjected to two identical testing paradigms to assess the reliability of the responses and directional tuning assessments. For each of these cells the optimal direction and tuning to non-optimal directions was highly similar across tests. Figure 5.6 gives an example of the similarity of results for one cell tested twice for directionality (with a delay of 5 minutes filled with other testing). In both tests a variety of different objects were moved by the experimenter under strong diffuse room lighting, including hands, gratings and other objects. If the directional selectivity were due to co-incidental variations in speed, position or form it would be expected that the cell responses would show markedly different tuning curves. As can be seen the similarity was remarkably high. This shows that the directionality estimates obtained here were accurate using hand held stimuli and quantitative analysis of cell response magnitudes. Indeed, other investigators have noted that even hand held testing with subjective assessment of responses can yield good estimates of directionality compared with computer controlled stimulus presentation, data collection and accurate fixation (Thier and Erickson 1992).

It is of course possible that speed sensitivity could have influenced directional tuning estimates. However directional testing was done at speeds within cell's optimal speed range. The testing revealed that most STPa cells were not selective for speed over the range tested. This was not due to the range of tested speeds being too small since the method was sensitive enough to pick up speed sensitivity amongst some STPa cells (see Figure 5.1). The consistency of

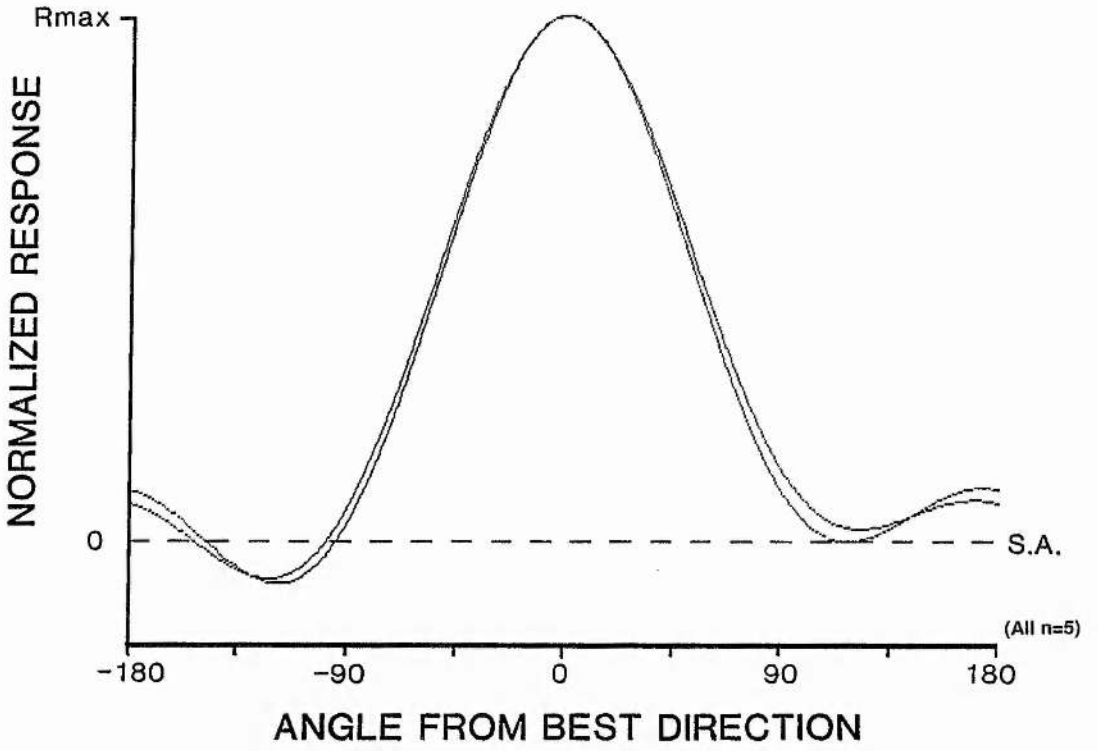


FIGURE 5.7. SIMILARITY OF TUNING ACROSS PLANES. Standardized regression curves of one cell sensitive to movement downwards to testing in two planes (sagittal and fronto-parallel, stimulus: experimenters hand subtending approximately 5 degrees moving at 30 degrees/s). The curves have been aligned to the peak response. The similarity between the two curves shows that the tuning for direction is almost identical in both planes. Regression: sagittal plane $R^2 = 0.648$, $F_{[4,35]} = 16.1$, $p < 0.0005$; fronto-parallel plane $R^2 = 0.768$, $F_{[4,35]} = 28.9$, $p < 0.0005$.

stimulus speed during testing was thus sufficient to pick up speed sensitivity amongst STPa cells and more importantly cannot account for the directional tuning estimates (see Figure 5.6).

Tuning across different planes

For 9 cells, directionality was studied in two orthogonal planes. Each of these cells was found to show highly similar directional tuning functions in the two planes. Figure 5.7 shows the result from one representative cell. One curve was derived from testing in the sagittal plane, where the estimated optimal direction was down and slightly towards the subject (14 degrees off vertical). The second curve represents tuning in the fronto-parallel plane where the optimal direction was down and slightly to the left (9 degrees off vertical). To facilitate comparison across planes, the tuning curves in the figure have been shifted so that the peaks are co-incident and the response magnitudes normalized (Rmax was 17.2 and 15.2 spikes per second in the two planes).

Discrimination between directions

In a study of the tuning of cell responses to different views of the static head and body a different discrimination index was used to quantify discrimination between different perspective views (Perrett et al. 1991). The index was defined as (R_{min} -- S.A.) / (R_{max} -- S.A.) where R_{min} and R_{max} were the minimum and maximum responses, respectively, to different views.

To facilitate comparison between tuning for perspective view and direction in the same brain area the same discrimination index was computed for the cells studied here. As ANOVA revealed no significant differences between planes for the index (p = 0.28) the data for cell testing in different planes were combined. Figure 5.8 shows the frequency histogram of the discrimination index for 55 cells. Only the first estimate is given for cells which were tested more than once. For all cells the response to the worst direction of motion was less than 1/2

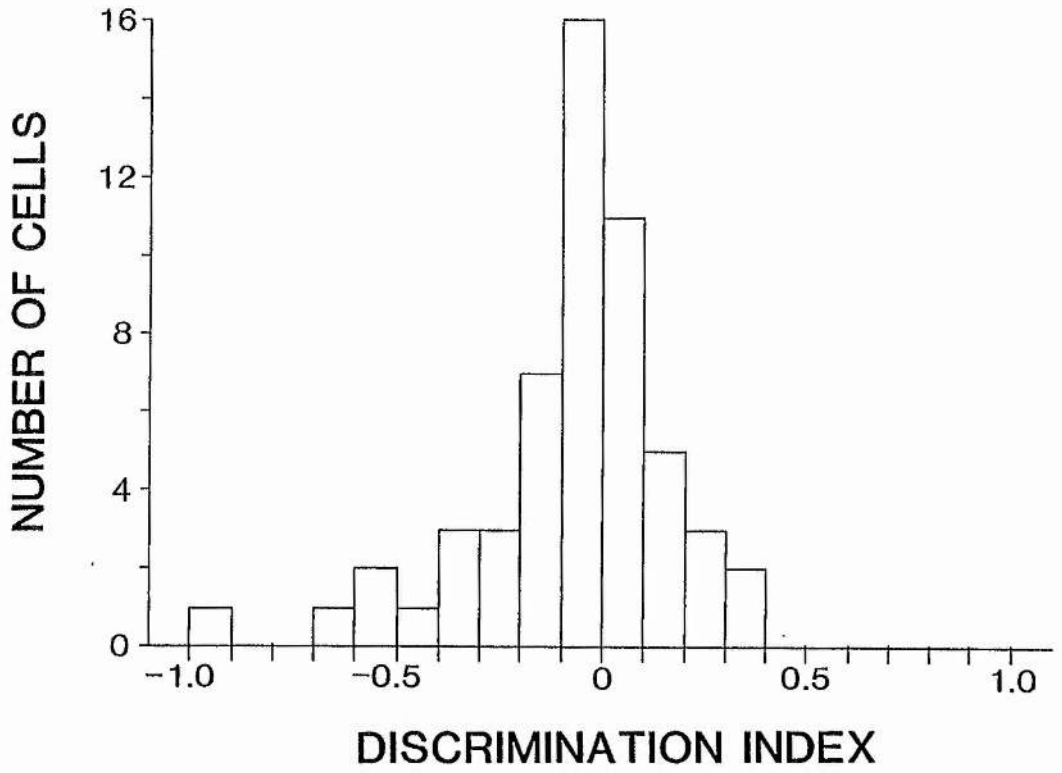


FIGURE 5.8. INDEX OF DISCRIMINATION: COMPARING BEST AND WORST DIRECTIONS. The responses to the least effective direction (Rmin) are expressed as a fraction (Rmin-S.A.)/(Rmax-S.A.) of the responses to the most effective direction (Rmax). S.A. = spontaneous activity.

of the response to the optimal direction. Negative index values indicate responses to non-preferred directions were less than the spontaneous activity.

For 55 directional selective cells the mean value of the discrimination index was --0.07 (+/-- 0.032 SE). The discrimination of direction amongst this population of motion sensitive cells was compared to discrimination of perspective view amongst cells selective for the static form of the head and body. The discrimination index measured for 110 view sensitive cells was found to have a mean of 0.04 (+/-- 0.031, Perrett et al. 1991). The discrimination of direction was significantly greater than the discrimination of perspective view ($t_{[144]}$ = 2.378, $p < 0.02$, using Satterthwaite's approximation for heterogeneity).

A number of other indices have been used to estimate the discrimination between directions. A commonly used direction index is $I_d = 1 - (R_{opp} - S.A.) / (R_{max} - S.A.)$, where R_{opp} is the response magnitude to motion in the opposite, or null, direction to that which gives the maximal response magnitude (R_{max}). For comparison with directional tuning in other studies, this index was also calculated (see Figure 5.9). Not surprisingly, bidirectional cells showed less discrimination between preferred and opposite directions than cells displaying unidirectional responses. Consequently bidirectional (Figure 5.9a) and unidirectional (Figure 5.9b) cells have been plotted separately. The mean value for the 44 unidirectional cells tested with stimulus motion in the null direction was 1.01 (+/-- 0.037).

Width of tuning of direction-selective cells

Width of tuning was calculated as the average angle required to reduce firing rate to half of the difference between response to the most and least effective directions [(Rmax -- Rmin)/2 or 1/2 width at 1/2 height measure]. As no significant differences were found for cells sensitive to motion in different planes ($p = 0.26$), the estimates from all 3 planes were combined.

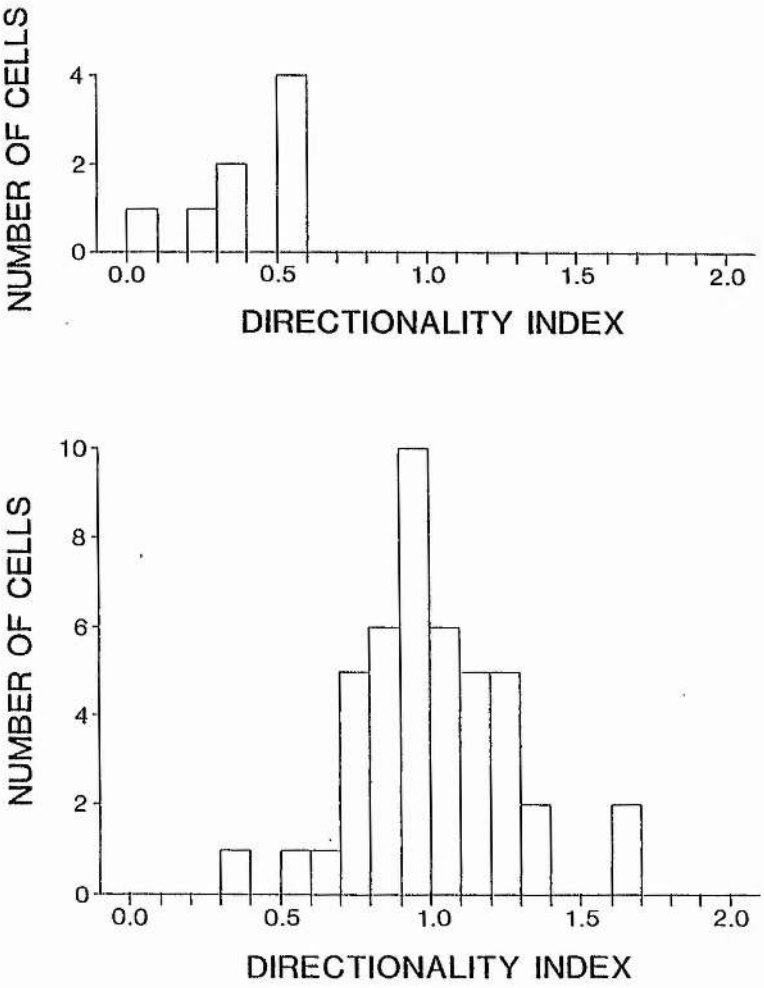


FIGURE 5.9. INDEX OF DIRECTIONALITY: COMPARING BEST AND OPPOSITE DIRECTIONS. Responses to direction 180 degrees from the optimal direction (Ropp) are expressed as a fraction of responses to the optimal direction (Rmax) [(Rmax-Ropp)/(Rmax-S.A.) or 1 - (Ropp - S.A.)/(Rmax - S.A.)]. S.A. = spontaneous activity. A directionality index of 0 indicates no difference in response magnitude to the two directions, a value of 1 indicates that there was no response above SA to the null or opposite direction and values > 1 indicate suppression of activity below S.A. to motion in the opposite direction. **UPPER:** cells classified as bidirectional. **LOWER:** cells classified as unidirectional.

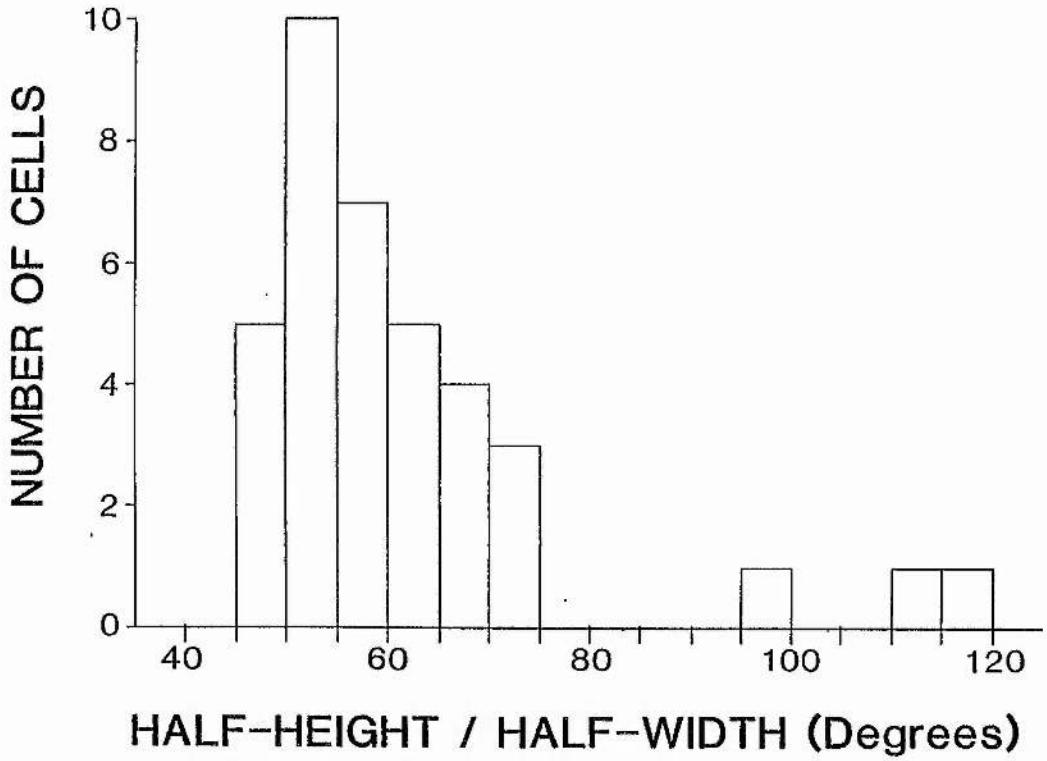


FIGURE 5.10. WIDTH OF DIRECTIONAL TUNING. The average angle of rotation required to reduce response by half of the difference between response to the most and least effective direction {(Rmax-Rmin)/2} is plotted for 37 direction selective cells.

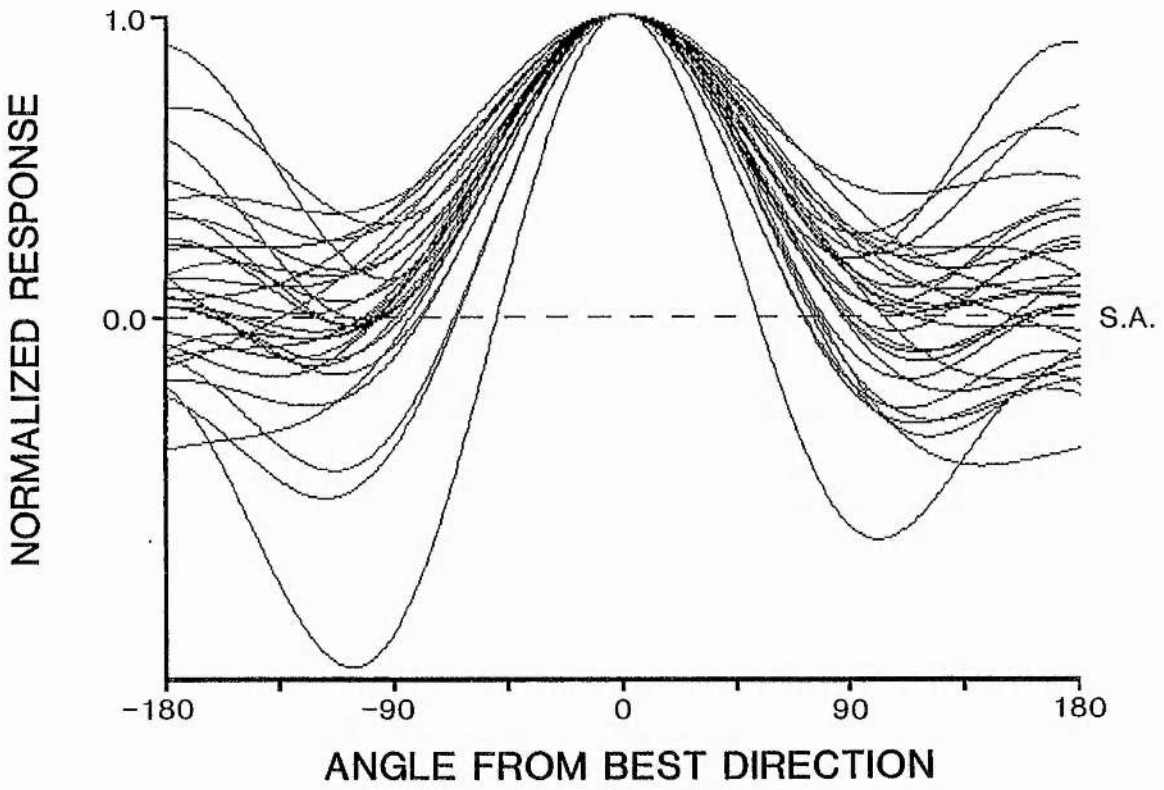


FIGURE 5.11. TUNING CURVES OF DIRECTION-SELECTIVE CELLS DISPLAYING UNIDIRECTIONAL NARROW TUNING. The tuning curves (estimated from best fit cardioid function relating response to angle of direction) for 29 unidirectional direction-selective cells. Each tuning curve is normalized so that maximum response = 1.0 and spontaneous activity (S.A.) = 0. Direction is expressed as an angle of rotation from optimal direction for each cell (qmax).

Figure 5.10 illustrates the width of tuning estimate for all direction-selective cells tested with 8 directions and with significant cardioid regressions (1 value per cell). Half width at half height ranged from 45° to 120°. The distribution of directional tuning in Figure 5.10 illustrates two important points: first the majority of cells have tuning less than 75° (1/2 height 1/2 width) and second, the distribution does not appear continuous. Using 90 degree 1/2 width tuning as a cut off point, 34 cells were defined here as having relatively `narrow' tuning for direction and 3 cells as having `broad' tuning

The distribution of width of tuning is skew positive with no cells having 1/2 height 1/2 width less than 45°. This is in part an artifact of regression analysis since the cardioid equation used cannot follow changes in response from maximum to minimum in less than 90°. From visual examination of the tuning curves and cell responses, for only 5 cells the estimated width of tuning was artificially broad (by an estimated 5--15 degrees). Since this error affected a minority of cells only, it does not affect estimates of the width of tuning of the cell population unduly. The distribution of values were compared with those obtained for cells sensitive to different views of head and body using the non-parametric Mann-Whitney U test and found to be similar (direction median 1/2 width = 55.5°, view median = 57.7°, $U_{[54,34]} = 1374.5$, $p = 0.24$).

Average shape of direction tuning

To make a visual comparison across different tuning curves, the raw data for each cell were rescaled (so that Rmax = 1.0 and S/A = 0.0) and directions expressed as angles of rotation from optimal (Soodak and Simpson 1988). Figure 5.11 illustrates the range of the individual tuning curves for a sample of 29 unidirectional cells with narrow tuning.

To obtain the average tuning curve for different cells, the coefficients of the regression analysis (Equation 1, above) of normalized data were averaged. Figure 5.12 displays the average tuning curves for uni and bi-directional cells. For

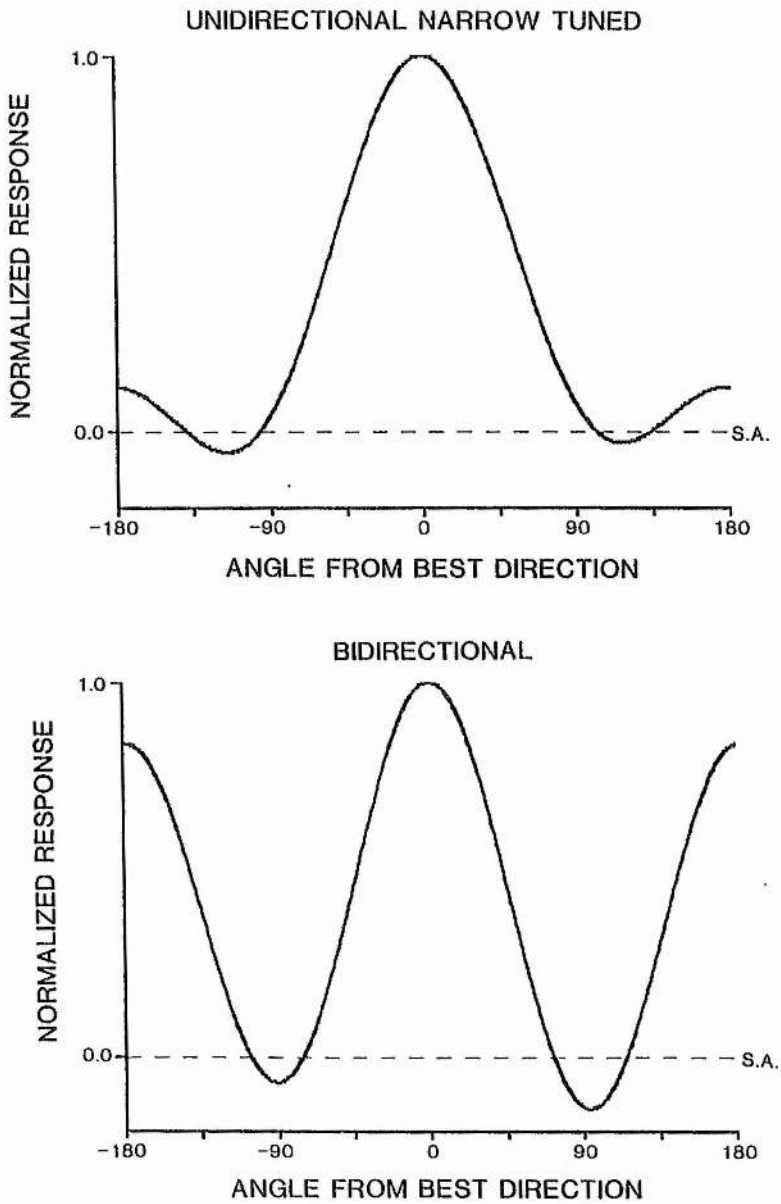


FIGURE 5.12. AVERAGE TUNING CURVES FOR DIFFERENT CLASSES OF CELL. UPPER: NARROW BAND DIRECTION-SELECTIVE CELLS (n=29). LOWER: BIDIRECTIONAL DIRECTION-SELECTIVE CELLS (n=5).

both types of cell the average tuning curve exhibits a dip in response to directions some 90° away from optimal view. This dip falls below spontaneous activity and may well arise from inhibition from cells tuned to these orthogonal directions.

Distribution of optimal directions

The optimal response directions were analysed for cells which (a) were tested with 8 directions, (b) displayed a significant ($p < 0.05$) relation between response and a cardioid function of angle of motion (Equation 1) and (c) for which Chi-Squared comparisons between predicted and observed response indicated a good fit of the cardioid function. Thus, data were considered for only those cells for which regression analyses produced appropriate optimal response angles.

Figure 5.13 shows the distribution of the optimal directions from the 46 appropriate analyses. For each cell the optimal direction is represented by a single line. For cells tested twice in the same plane the first estimate of optimal direction is plotted. Where testing was performed in two planes both estimates have been entered in the appropriate figure.

As shown in the upper part of Figure 5.13, the optimal directions of cells appear clustered around cartesian axes, (up/down, left/right and towards/away). To evaluate this clustering the estimated optimal direction is expressed as the angular rotation from the nearest cartesian axis (Figure 5.13 lower). Statistical analysis confirms that significantly more cells have optimal directions that are 'on axis' (within 22.5 degrees to a cartesian axis) than would be expected by chance (Binomial Test $p < 0.0005$).

Temporal characteristics of cell responses

The temporal characteristics of the responses of cells within STPa to static form were investigated (Oram and Perrett 1992; see chapter 3). Whilst there were insufficient data available to perform a complete analysis of the same

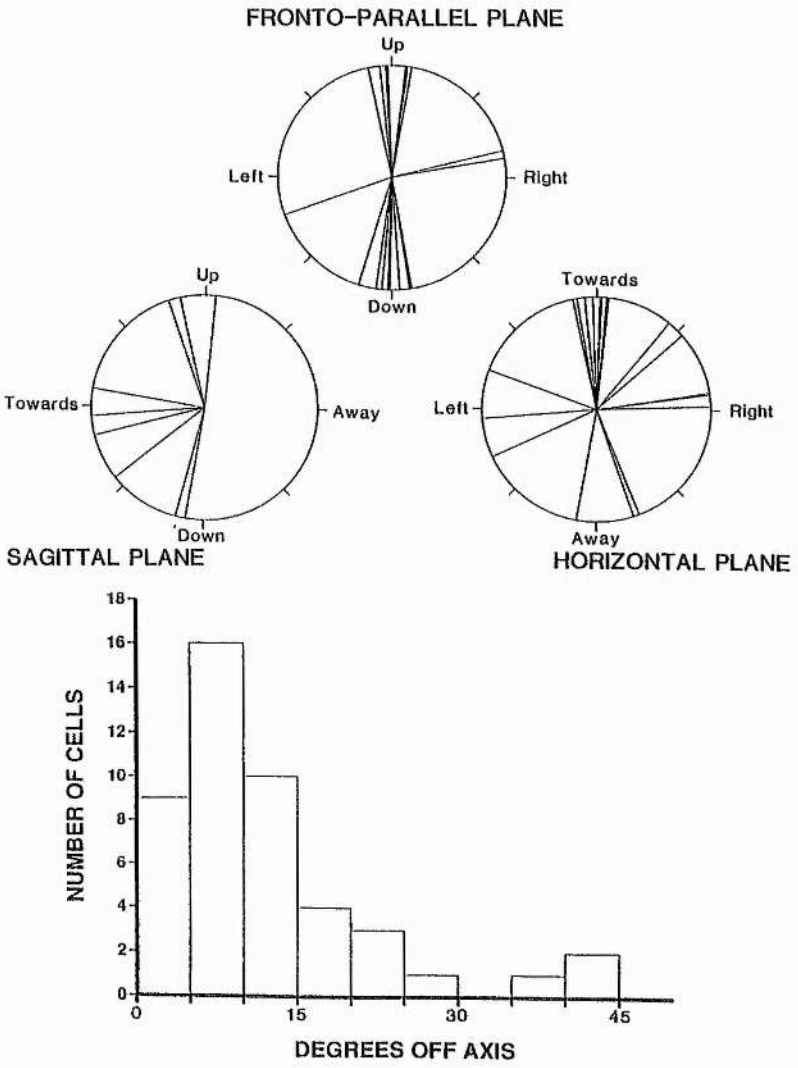


FIGURE 5.13. THE DISTRIBUTION OF PREFERRED DIRECTIONS ACROSS THE POPULATION OF MOTION SENSITIVE STPA CELLS. **UPPER:** polar plots, with each line representing the direction estimated from regression analysis to evoke maximal response for one cell, plotted separately for each plane. **LOWER:** histogram of the number of cells for a given rotation away from a cartesian axis independent of the plane of testing. Significantly more cells exhibit a preference for directions within 22.5^{o} of cardinal directions (up, down, left, right, towards and away) than for intermediate directions (44 out of 46, $p < 0.0005$).

characteristics of cells selective for the direction of motion in the same brain area, analysis allows a partial comparison of the two populations.

Table 5.2

	Direction		Static Form*
	Mean	Range	Mean
Timing (ms)			
Latency	90.9	35.0 -- 126.4	119.1
Rise time	69.4	10.4 -- 175.0	58.2
Half fall time	59.0	20.0 -- 124.8	40.0
Decay time	134.4	20.8 -- 754.0	93.4
Duration	168.6	15.6 -- 764.4	112.5
Firing rates (spikes/s)			
S.A.	11.4	0.8 -- 40.8	8.6
Peak	108.3	62.2 -- 175.9	115.1
First 100 ms	67.3	34.7 -- 89.2	66.9
Second 100 ms	53.1	20.7 -- 84.0	48.1
Fifth 100 ms	31.9	3.0 -- 56.3	28.5
End 100 ms	30.6	5.4 -- 59.2	24.7

Table 5.2. Time-course of responses of motion sensitive STPa cells lacking form selectivity. * Measurements from STPa cells involved in static form processing and selective for the perspective view of the head for comparison (from Oram and Perrett 1992).

Latency estimates were obtained for 15 cells where the data had been collected to 5.0 or 5.2 ms accuracy. The mean was 90.9 ms (See Table 5.2). The static form cells, under similar presentation conditions had a mean latency of 119 ms (Oram and Perrett 1992; see chapter 3). These are statistically different ($t_{[57]} = 3.34$, $p = 0.001$), with direction selective cells responding at earlier latencies than form selective cells. As response magnitudes were almost identical (67.3 and

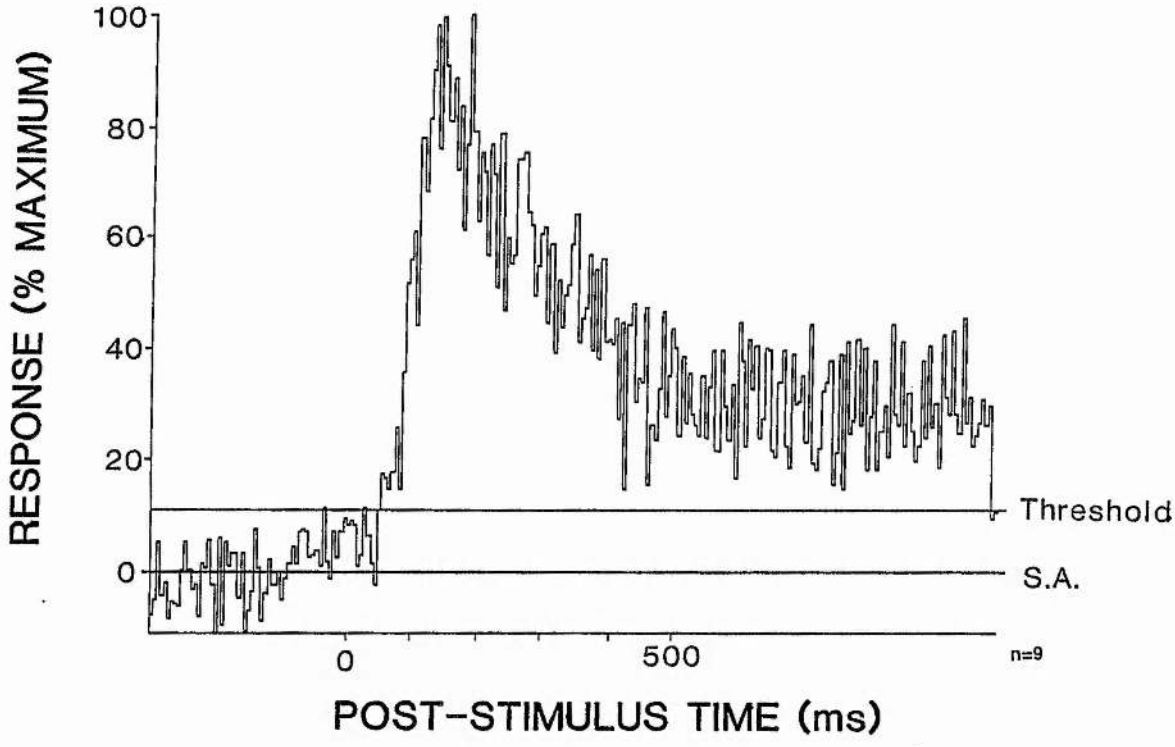


FIGURE 5.14. POPULATION RESPONSE TO OBJECTS MOVING IN THE MOST EFFECTIVE DIRECTION. The combined responses of 9 cells are plotted as a percentage of the peak response. The population response latency was estimated at 58.8 ms. S.A. = spontaneous activity.

66.9 spikes/s during the first 100 ms of the response), this difference in latency was not due to stronger responses in the cells of the present study.

As with cell responses to static form information, the responses of the cells in the present study to moving objects showed a fast rising phase to a peak, then a more gradual decay down to an apparently steady firing rate. The steady firing rate (estimated from the final 100 ms of the data collection period) was found to be above the spontaneous activity level for all cells.

Table 5.2 summarizes the parameters used to define the time course of responses for 15 motion sensitive cells lacking form selectivity. The mean values for STPa cells selective for static form (different views of the head) are given for comparison. The Table indicates the overall similarity of the temporal profile of the responses of the two cell types.

Figure 5.14 shows the population average of the responses of 9 cells where spike activity was collected in 5.2 ms time bins. The population latency of this sub-sample was 58.8 ms. This is shorter than the population latency for cells sensitive to static form (95 ms, Oram and Perrett 1992; see chapter 3), confirming the statistical assessment of the individual latency estimates. Other temporal measures of the population response of direction sensitive cells were similar to those obtained for the view selective population response (e.g. rise times of 69.4 and 62.4 ms for direction selective and view selective population responses, respectively).

Location of cells

Histological reconstruction of the positions of cells recorded in monkeys F,B,D (e.g. Figure 5.15) indicated that the majority of non-form motion sensitive cells were located in the cortex of the upper bank of the superior temporal sulcus (areas TPO and PGa of Seltzer and Pandya 1978). The proportions of cells responsive to movement but lacking form sensitivity out of the total number recorded within STPa varied from subject to subject (B, 10.6% (67/632); D,

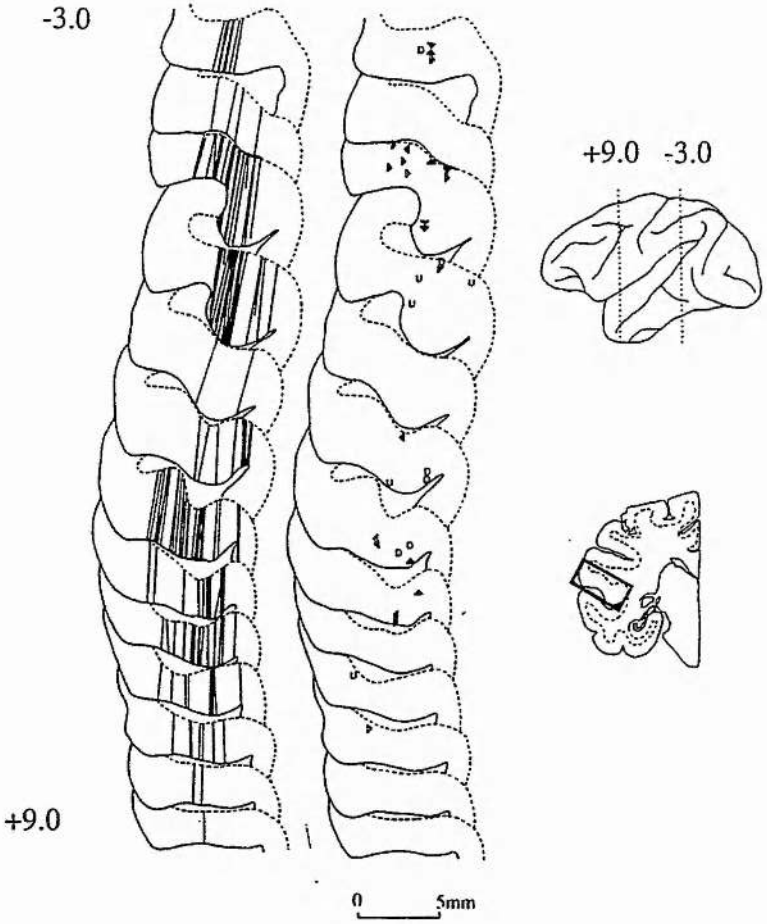


FIGURE 5.15. HISTOLOGICAL RECONSTRUCTION OF POSITION OF CELLS RESPONSIVE TO MOTION BUT LACKING FORM SENSITIVITY. UPPER RIGHT: lateral view of the right hemisphere showing the antero-posterior extent of sampling. LOWER RIGHT: a coronal section at 3.0 mm posterior to the mid-geniculate level. The box shows the position of the superior temporal sulcus (STS). LEFT: serial sections (every 1.0 mm) of the upper bank of the STS from 3.0 mm posterior to 9.0 mm anterior to the mid-geniculate level from the right hemisphere of one monkey (D). Vertical lines indicate the position of all recording tracks in this hemisphere. CENTRE: position of movement sensitive cells. The preferred direction of motion for each identified cell is indicated by the symbols: up and down arrows (motion towards and away), left and right arrows (motion left and right), U and D (motion up and down).

11.1% (155/1397); F, 14.8% (225/1553). (N.B. these figures include cells responsive to motion that were not investigated in detail for direction selectivity.) The motion sensitive cells constituted approximately 25% of all visual cells in the STPa. Measurements of the position of recording electrodes (from X-radiographs) indicated that cells sensitive to motion but not form in monkeys J and H were recorded in the same region.

Figure 5.15 illustrates the reconstruction of the position of directionally selective cells that were recorded in the upper bank of the STS in the right hemisphere of one monkey (D). Neighbouring cells on the same track showed a tendency to display similar direction preferences, though within a given 1.0 mm patch of cortex multiple directions appear to be encoded. Thus with the resolution of reconstruction present (+/-- 1.0 mm) there was no obvious anatomical organization of direction coding within the cortex of this monkey at a macroscopic level.

DISCUSSION

Motion-coding in temporal cortex

Previous work with cells responsive to visual stimuli in the STPa focused mainly on coding of information about the form of the stimulus. Indeed, the cortex of the temporal lobe is often considered as containing high level representations of form. The current study evaluated coding of direction information by cells with no apparent sensitivity to the form of the stimulus. Few studies have investigated motion processing within the temporal cortex, yet there is evidence that motion can be used as a source of information to define form within this area (Britten et al. 1992; Oram and Perrett 1994a,b; Perrett et al. 1990a, 1990b; see following chapter). The motion sensitive cells studied here were found in the same locus (within the upper bank of the anterior superior temporal sulcus or STP) as cells selective for the static form of the head and body

(Perrett et al. 1991). This co-localization emphasises the anatomical convergence of streams of information processing for form and motion.

Coding of direction in STPa

In keeping with the relatively low rate of spontaneous activity, *optimal* directions appeared to be coded by excitation rather than inhibition. That is, no cells were found coding the presence of one direction by a selective reduction of response rate below spontaneous activity with no change in activity for other directions of motion. There was, however, some evidence for a role of inhibition in the coding of non-preferred directions.

It is interesting to note the `Mexican hat' shape of direction tuning amongst the majority of directionally selective cells. The dip in response, for directions approximately 90^{o} to the cells' preferred direction, may represent inhibitory interactions between cells tuned to different directions. Indeed inhibition relative to spontaneous activity was observed for many cells to *non-optimal* directions and also static stimuli (e.g. Figure 5.1). Inhibition has been reported to motion in the null direction for some 90% of MT cells. Direction selectivity in MT would seem to be established by both inhibition in the null direction (as suggested by Barlow and Levick 1965) and by facilitation of the response in the preferred direction (Mikami et al. 1986a). The degree of inhibition for non-optimal directions exhibited by STP cells was not as great as that reported for cells in other visual systems (e.g. the rabbit accessory optic system, Soodak & Simpson 1988).

The amount of inhibition found in the present study of STPa direction coding was substantially greater than that found in STPa coding of the perspective view of static objects. Approximately half (14/29) of the tuning curves for individual directionally tuned STPa cells drop numerically below spontaneous activity (Figure 5.11). For cells tuned to different views of the head, only one quarter (11/43) of the individual tuning curves drop below spontaneous activity

levels (see Figure 8, Perrett et al. 1991). The difference in the amount of suppression seen in these two cell populations is reflected most clearly by comparison of the average tuning curves. In the present study the average tuning curve for direction drops to slightly below spontaneous activity for non-optimal directions (Figure 5.12a), whereas the average view tuning curve for cells responsive to the head remains well above spontaneous activity throughout the full 360 degrees of head rotation (see Figure 9a of Perrett et al. 1991). In a detailed examination of the responses of cells selective to static head views, very little evidence was found of inhibition to non-optimal views (Oram and Perrett 1992; see chapter 3). The increased suppression and/or inhibition seen in motion selective STP cells compared with form sensitive STP cells may therefore reflect a qualitatively different process for establishing tuning.

The range of the direction tuning defined by 1/2 height 1/2 width measure in the present study was 45 to 120 degrees. This range was similar to the range of tuning exhibited by cells selective for head view. The difference in the amount of inhibition seen between form and motion sensitive cell responses does not therefore lead to tighter tuning of motion sensitive cell responses.

Directional tuning and cartesian axes

Previous work has emphasized the prevalence of viewpoint sensitive coding for static form information in the STP (Bruce et al. 1981; Desimone et al. 1984; Perrett et al. 1982, 1985a, 1991, 1992; Kendrick and Baldwin 1987; Hasselmo et al. 1989). The current study also indicates that motion processing in temporal cortex is heavily influenced by the observer's viewpoint. Of course, in most brain areas direction of movement is specified relative to the viewer. Some processing of motion in the temporal cortex, however, is conducted in an object-centred framework where the direction of object movement can be understood best when specified relative to parts of the object being viewed. For example head moving to chest regardless of orientation of the body relative to the viewer (head-

nodding, Hasselmo et al 1989) or arm moving to chest (Perrett et al., 1990), a person walking forwards (following their nose) as opposed to backwards (Perrett et al., 1985; Oram et al., in submission; see chapters 7-9). In the present study it was found that half of the motion sensitive cells in STPa were selective for particular directions of motion relative to the observer. This viewpoint sensitivity could allow the motion information to be combined more easily with the form information that is processed in the same brain area.

One of the clearest findings of the present study was the prevalence of direction-selective coding clustered around particular cartesian axes. These axes correspond to the gravitational axis (up/down), an axis running towards/away along the line of sight and an axis running left/right. The width of the tuning of the cells suggests that coding of these 6 directions is sufficient to allow representation of all possible directions of movement within the STP. Movement at 45 degrees to the cartesian axes would excite (half maximally) at least two cell populations tuned to directions along the cartesian axes. The preferential coding of orthogonal directions means that all directions in three dimensional space can be represented by the minimum number of directionally selective cell populations. Therefore, direction of motion is represented in STPa in the most efficient manner. Although an initial investigations of V5 (MT) suggested a bias in cell response preferences for movement towards the contralateral periphery (Dubner and Zeki 1971), subsequent studies have not drawn attention to any pronounced bias in the distribution of directional tuning (Zeki 1974; Albright 1984 and see introduction). Cells in MST also exhibit no marked bias in directional preference, except in the horizontal plane where cells show a slight preference for motion towards the ipsilateral side (Komatsu and Wurtz 1988a), a property also seen in the visual tracking neurons in the lateral part of MST, MSTl (Thier and Erickson 1992). The biases reported for MT and MST were towards a particular hemispace (i.e. contralateral and ipsilateral) and not the strong preference along particular axes reported here.

Preferential coding of direction has been noted in sub-cortical structures. The neurons of the accessory optic system (AOS) of the rabbit, cat and monkey show preferential coding of directions of movement that coincide with the direction of retinal movement that would occur during self motion about the vestibular axes (Soodak and Simpson 1988; Simpson et al. 1988; Grasse and Cynader 1982, 1984; Grasse et al. 1984; Hoffman and Distler 1989). Coupling the sensory inputs from the semicircular canals with the optical changes resultant from self motion could also be used to define the three axes of motion selectivity observed in STPa. Indeed real motion and retinal movement consequent on self motion have markedly different effects on STPa cells selective for motion (Hietanen and Perrett 1992).

Given the importance of the cartesian axis system for the coding of motion and view it is also interesting to speculate on the development of preferential coding. The direction of motion of an animate object will be highly correlated with the perspective view of the moving body. For example, the left profile head view would be correlated with motion of an animal to the observer's left. This relationship could underlie the correlation between the views preferentially encoded in STPa (front, left and right profile and back) and the directions preferentially coded in the same area (towards, left, right and away). Furthermore, the cells coding head and body information associated with another individual's attention up and down could also be related to the body movements up and down (Perrett et al. 1992).

It is of interest to consider why these particular three axes are represented. If directional tuning is affected by experience then the axes utilized in STPa are not too surprising. The up/down axis is, of course, coincident with gravity; this axis could be defined through experience of objects falling. Movement towards any organism has strong survival implications and, for social animals such as macaques and humans, it is a powerful cue to social interactions. From optical considerations objects moving along this z axis will change in retinal size. The

presence of retinal expansion/contraction could therefore be used to define the selectivity of cells tuned for movement towards and away. With two axes defined, the third axis (left/right) can also be found by a system of coding which attempts to decompose movement into uncorrelated or orthogonal directions. Learning rules which maximise the difference between activity amongst cell populations (e.g. 'decorrelation' of Foldiak 1991) will automatically extract orthogonal dimensions (or principal components) amongst sensory inputs.

Relationship of motion processing in STPa to posterior areas

The obvious route for motion information to arrive in area STPa is from the motion pathway involving the magnocellular portions of the lateral geniculate nucleus (LGN) and areas V1, V2, MT and MST. Motion sensitive visual inputs to STP may not be entirely dependent on the magnocellular-geniculostriate system. The visual responses in MT are largely dependent on the magnocellular pathway (Maunsell et al. 1990) but, after lesion of magnocellular LGN, there is evidence for additional parvocellular input to MT (Merigan et al. 1991). Motion sensitivity in STPa could also depend on inputs from the superior colliculus, which provides inputs to extrastriate visual areas via the lateral pulvinar (Bruce et al. 1986; Girard and Bullier 1989; Girard et al. 1991; Gross 1991; Rodman et al. 1989, 1990). The colliculus also projects to S and interlaminar layers of the LGN. These layers are the source of direct projections from the LGN to extrastriate areas V2, V3, V3a, V4 and MT (Benevento and Yoshida 1981; Bullier and Kennedy 1983; Fries 1981; Wong-Riley 1976), though there do not appear to be direct LGN connections to anterior temporal cortex (including area STP) or parietal cortex (area PG) (Fries 1981; Iwai et al. 1980; Yukie and Iwai 1981).

It has been proposed that the residual vision seen after lesions to striate cortex is mainly due to neuronal activity in the dorsal pathway (Girard et al. 1991). Following unilateral lesions to striate cortex many STP cells remain

visually responsive but selectivity for form and motion direction is largely abolished (Bruce et al. 1986; Gross 1991). Bruce et al. (1986) found only two out of forty STP cells exhibiting residual directional selectivity. Thus, although the superior colliculus may provide some visual input to the STP, it does not seem to be the main source of directional selectivity in this area.

Interestingly, directional selectivity is maintained in many MT cells following V1 lesion, although response magnitudes are reduced (Rodman et al. 1989). This residual directionality is presumably due to input to MT from the superior colliculus, since lesions of both striate cortex and the superior colliculus abolished all visual responses in MT and STP (Rodman et al. 1990; Bruce et al. 1986). It is not clear why the residual sensitivity to motion in MT which survives striate lesion is insufficient to drive directional selectivity in more than a minority (5%, 2/41) of STP cells (Bruce et al., 1986).

Temporal characteristics of the response

Responses of MST cells show a phasic increase from response onset which lasts for some 100--300 ms. The tonic response which occurs after this transient burst lasts for the duration of the stimulus presentation (Duffy and Wurtz 1991a). This pattern of a transient burst followed by a period of tonic discharge throughout stimulus presentation was also seen in both STPa cells sensitive to motion (Figure 5.14) and cells responsive to static stimuli (Oram and Perrett 1992; see chapter 3). The latency of MT neurons has a mean of 58 ms, whereas the latency of the population curve of direction selective STP neurons is 68 ms (Figure 5.14). This is consistent with motion information arriving in STPa via MT. It has been argued elsewhere that static information in STPa neurons arrives as quickly as possible via the ventral route (Oram and Perrett 1992; see also chapter 3), yet the latency of the cells in the present study was shorter than the latency of cells responding to static form. Most likely therefore, the motion input derives from the dorsal route (involving MT and MST) as there is a pathway from

V1 to STPa passing through fewer brain areas than in the ventral route. Conversely the static pattern input would derive from the ventral route involving V4 and inferior temporal cortex (Young 1992; Felleman and Van Essen 1991).

Proportions of motion sensitive cells

In areas MT, MST and posterior STP (STPp), virtually all visual cells are non-selective for stimulus form but responsive to motion (98% of MT, 96% MST, Tanaka et al. 1986; MST 99%, STPp 97% Hikosaka et al. 1988; MSTd and Lateral MST, MSTl 96% Saito et al. 1986; MSTd 85%, Duffy and Wurtz 1991a). This is far above the proportion (25%) of visual cells in STPa displaying motion sensitivity but lacking form sensitivity.

A further difference between non-form, motion selective cells in STPa and those of MT/MST lies in the increased number of pandirectional cells. Studies of MT revealed only 2% of motion selective cells were pandirectional (Tanaka et al. 1986). A similarly low proportion (1%) of pandirectional cells was found in MSTd and MSTl (Tanaka et al. 1986, Saito et al. 1986). In a third study of MSTd, some 30% of the cells were responsive to rotation, movement in the fronto-parallel plane and expansion/contraction but, of those tested in the fronto-parallel plane, only 7% responded to all 8 directions tested (Duffy and Wurtz 1991a). All these estimates are far lower than the 56% of STPa motion sensitive cells found in the present study to be pandirectional. Hikosaka et al. (1988) also report that 30% of motion sensitive cells are pandirectional in STPp. Thus, the change in the ratio of unidirectional to pandirectional cell selectivities seen from MT/MST to STPa is due to a drop in the number of unidirectional motion selective cells, 86% in MT and 80--88% in MST (Tanaka et al. 1986; Saito et al. 1986). The proportion of bidirectional cells remains consistent across areas MT (6%, Tanaka et al. 1986), MST (1%, Saito et al. 1986) and STPp (4%, Hikosaka et al. 1988) and is far more comparable to the 4% found in STPa in the present study.

Of the unidirectional cells in posterior regions relatively few respond to expansion or contraction; MT, 3%; MST, 5--23%; STPp, 13% (Tanaka et al. 1986; Saito et al. 1986; Hikosaka et al. 1988). In the present study 46% of unidirectional cells responded to movement towards or away from the monkey. A similarly high number of cells responding to movement along the z axis were found in STPa polymodal cells (Mistlin and Perrett 1990). This apparent discrepancy between posterior areas and STPa cell populations could be due to the observation that expansion/contraction cells respond preferentially to real 3-D stimuli over projected stimuli (Tanaka and Saito 1989; cf Hikosaka et al. 1988). The ratio of unidirectional cells in MSTd responsive to expansion compared to contraction ranges from about 2:1 (Saito et al. 1986) to 7:1 (Tanaka and Saito 1989). The ratio of the present study was 67:28. Thus in all areas there are more cells responsive to motion towards than away from the monkey but the proportion of cells sensitive to motion along the z axis is much higher in STPa than in MT, MST or STPp.

Thus the changes that seem to occur between posterior motion areas (MT, MSTd/MSTl, STPp) and STPa can be summarized as follows: (1) the proportion of motion sensitive cells decreases; (2) there is an trend for fewer unidirectional cells and more pandirectional cells; (3) the proportion of bidirectional cells stays roughly equivalent indicating that they form a separate population and (4) the proportion of motion sensitive cells preferring motion along the z axis (towards and away) increases.

Directional tuning in different brain areas

It is also of interest to compare the direction index measures (Id, Figure 5.9) with other extrastriate brain areas. A number of measures of Id have been made for MT (e.g. median = 0.99, Saito et al. 1989; 1.01, Snowden et al. 1992). Furthermore, the estimate of Id is similar when isoluminant, colour contrast stimuli are used (median 0.95 Saito et al. 1989). The median Id of the

unidirectional STPa cells in the present study was 0.97 (Figure 5.9b). In another study, more than 80% of the unidirectional cells in MT had Id values greater than 0.8 (Mikami et al. 1986a). As the criteria for including a cell in their analysis was very stringent, the estimate is very unlikely to include any bidirectional cells. For the unidirectional cells in the present study, it was also found that greater than 80% of cells had Id values exceeding 0.8.

Estimates of directionality in MST (Duffy & Wurtz 1991a) are similar to those obtained here, though a greater proportion of STPa cells appear to have higher direction discrimination. Such a regional discrepancy may be due to the use of the minimum estimated response, whereas Duffy and Wurtz used the observed response magnitude to the null direction. Thus, the directionality in area STPa appears similar to that in MT and MST but greater than that estimated for V1 directional cells (e.g. I_d = 0.44, Snowden et al. 1992). This suggests that STPa, MT and MST code movement with similar discrimination between directions.

It is worth noting that the discrimination between directions shown by cells in STPa (Figure 5.8) is comparable with the equivalent measure for discrimination between views for cells selective for different views of the head in the same brain region (Perrett et al. 1991). Thus the discrimination, the breadth of tuning and distribution of preferred directions shown by STPa neurons processing motion appear comparable to that shown by neurons in the same area that process the static form of the head and body.

Establishment of direction tuning

It has been argued that the tuning seen in static form cells of STPa is derived in a feed forward way with little lateral inhibition contributing to the initial view selectivity. This was based on the observation that view discrimination was present from response onset and that response latencies were

so early that, even following the shortest route (involving only 8 synapses) along the ventral ('form' or 'what') pathway to the STPa, the responses had to be relayed between stages on the basis of the first one or possibly two spikes of the response (Oram and Perrett 1992; see chapter 3).

The latency of the cells in the present study is on average statistically less than the latencies of cells selective for views of the head and body. As firing rates during the initial 100 ms of response were comparable (67.3 as opposed to 66.9 spikes/sec) and similar stimulus presentation methods were used, this difference is unlikely to be due differences in the adequacy of the stimulus. The difference in latency suggests that the input to motion selective cells in STPa is established by a different route not involving the ventral pathway. The shortest route to STP from the striate cortex is via the dorsal pathway (V1-MT-MST/FST-STPa, Felleman and Van Essen 1991). Allowing for interneurons between input and output layers in MT and MST/FST (Felleman and Van Essen 1991; Zeki and Shipp 1988), a minimum of 5 synaptic relays are needed to pass information from V1 to the STPa cells recorded here. Latencies of V1 cell responses to strong contrast stimuli can be as early as 30 ms (Maunsell, personal communication). As the population latency of the STPa cells responsive to motion studied here was found to be 58.8 ms, this leaves 30 ms for activity to flow between V1 and STPa. Allowing for 4--5 ms for each synaptic relay, the difference in latencies fits well with a route through MST to the recording sites in STPa. The time constraints along this pathway seem therefore as tight as those for static form processing, suggesting again that information must be passed between stages on the basis of the initial spike of the response.

Several studies have suggested that directionality can be established in a feed forward manner (e.g. Worgotter & Holt 1991; Soodak & Simpson 1988), which is a requirement to obtain the fastest flow of information. Indeed, the changes in the creation of new directional selective properties in MST can be

explained by direct feed-forward input from MT (Saito et al. 1986. Tanaka and Saito 1989, Tanaka et al. 1989).

Unlike the cell population studied here, the optimal direction for MT and MST neurons is more evenly distributed around the fronto-parallel plane (Albright et al. 1984; Komatsu and Wurtz 1988a). From Figure 5.13 it can be seen that more than 75% of the cells have optimal directions within 15 degrees from the cartesian axes. If the directional tuning depended on input from MT/MST then the preferential coding of directions in STPa could be established from outputs of just a small fraction (5--10%) of MT/MST neurons.

Relationship to psychophysical studies

Motion processing at the cellular level in V1, MT and MST has been correlated frequently with motion sensitivity at the psychophysical level (Snowden et al. 1991; Mikami et al. 1986a; Movshon et al. 1985; Newsome et al. 1986). Using moving random dot displays, Levinson and Sekuler (1980) measured the elevation in luminance detection threshold for various directions following adaptation to motion of the display in one direction. The elevation in threshold varied with the direction of drift of the test stimuli from the adaptation direction. It was found that maximal elevation occurred when the test stimuli moved in the same direction as the adaptation stimulus and fell to a minimum for opposite directions of motion. From this the estimated band width of underlying directionally selective channels (1/2 width at 1/2 height) would be between 40 and 60 degrees, which is equivalent with the tuning (1/2 width at 1/2 height) of STPa unidirectional cells reported here.

Adelson and Movshon (1982, see also Movshon et al. 1985) suggested a two stage model of motion detection, with the first phase extracting local orientation based velocities and the second phase combining these to get a measure of whole field movement. Psychophysical studies with plaid patterns support this two stage model (Welch 1989) although the apparent direction acuity

of plaid displays is greater along the horizontal and vertical directions than in oblique directions (Heeley & Buchanan-Smith 1992). Thus, if there was a neuronal correlate to the psychophysical observations, the cell responses would be expected to show preferential tuning to horizontal and vertical directions, have large receptive fields and have a band width of tuning of 40--60 degrees. These properties, in particular the meridional anisotropy, correlate more closely with STPa cells than cells in MT/MST. Further work with plaid patterns would be needed to establish whether the properties of STPa cells account for the psychophysical results.

Conclusion

Measurements of response latency indicate that movement information arrives in the STPa via a different route from that supplying sensitivity to the form of biologically meaningful objects. The directional selectivity of cells within STPa is consistent with an input of motion information from posterior motion processing areas (MT and or MST). Unlike the posterior areas, however, the distribution of preferred directions amongst STPa motion sensitive cells is highly clustered around the cartesian axes (towards/away and left right, up/down). Cells in STPa that are selective for static information about heads and bodies also exhibit a preference for views clustered around cartesian axes (front and back, left and right sides). The employment of the same system of axes for visual processing by these two distinct cell populations may underlie the production of a further population of STPa cells which are conjointly sensitive to form and motion. The specialization of STPa cells for particular directions and object views may facilitate the integration of two streams of visual analysis and thus support the unified experience of moving forms.

CHAPTER 6

MOTION PROCESSING IN STPa `BIOLOGICAL MOTION' SENSITIVITY

(Oram & Perrett 1994, J. Cog. Neurosci., 6:99-116)

(Oram & Perrett 1994, Proc. SPIE 2054, 155-165)

INTRODUCTION

In the early 1970's it was found that human subjects could interpret extremely impoverished images of human walking. Small light sources were attached to the points of articulation of a walking person (the shoulders, elbows, wrists, hips, knees and ankles), then all other visual information removed by presenting the stimulus in darkness. Johansson (1973) found that subjects had no difficulty in identifying the stimulus as representing a person walking. Indeed Johansson reported the effect as being "immediate and compelling". He referred to this type of stimuli as biological motion stimuli. They have also been referred to as moving light displays.

Human subjects can perceive a variety of information from such biological motion stimuli, including the gender and identity of familiar individuals (Cutting 1978; Cutting & Kozlowski 1977; Cutting et al., 1978; Cutting et al., 1988; Kozlowski & Cutting 1977), whether the individual walks forwards or backwards (Perrett et al., 1990a,b; Mather et al., 1992), as well as the mode of ambulation (Jansson & Johansson 1973; Fox & McDaniel 1982; Bertenthal et al., 1985) and other actions (e.g. sign language, Poizner et al., 1981). Thus at the human perceptual level biological motion stimuli can give a great deal of information,

not only about the nature of the movements but also the form of the individual that is moving. Despite the rich source of information from such non-rigid motion stimuli little is known about the underlying neuronal mechanisms. The similarity of behavioural performance between human and macaque subjects in processing form from motion (Siegel & Andersen 1990) suggests that the macaque is a suitable model for investigating the underlying neural mechanisms of form from motion processing.

Form and Motion Pathways

A brief recap of the two major visual processing pathways is of relevance here. As noted in the introduction, it has been suggested that processing of visual information in primates follows two pathways: the ventral "form" or "what" pathway and the dorsal "motion" or "where" pathway (Ungerleider & Mishkin 1982; Mishkin et al., 1983; De Yoe & Van Essen 1988). These two pathways involve several brain areas (Felleman & Van Essen 1991; Young 1992). The ventral pathway passes through the areas V1, V2, V4, into posterior, central and anterior inferotemporal cortex (PIT, CIT, AIT) and the anterior sections of superior temporal sulcus (including area STPa). The dorsal or "motion" pathway flows from V1 through V2, the middle temporal area (MT), also known as V5, and the lateral and dorsal medial superior temporal areas (MSTl and MSTd) and then passes to the frontal eye fields and parietal cortex. The termination areas of this pathway have led to the suggestion that it is involved in the control of eye movements and visuo-motor interactions with objects (Goodale & Milner 1992). The two pathways are not completely separate: outputs from areas MSTl and MSTd also pass through the fundus of the superior temporal sulcus (FST) to the posterior and anterior sections of the superior temporal polysensory area (STPp and STPa, Boussaoud et al. 1990). Area STPa therefore receives inputs from both the ventral (form) and dorsal (motion) pathways (Felleman & van Essen 1991; Young 1992). In view of this anatomical convergence, it may not be surprising

that some neurons in area STPa show selectivity both for the form and the direction of motion of objects. Single cells in macaque STPa (and more generally throughout the anterior sections of the superior temporal sulcus, STS) have been found to be selectively responsive to the sight of various body movements including walking and articulation of individual limbs (Brothers & King 1992; Desimone et al., 1984; Hasselmo et al., 1989; Perrett et al., 1985b, 1990b; see also later chapters) and hand actions (e.g. tearing, object manipulation, Perrett et al., 1989a,b). In this section, the responses of cells to whole body motion defined under biological motion are described.

Mechanisms of sensitivity to form and motion

There are three broad possible categories of mechanism by which conjoint selectivity to form and motion could be achieved in the STPa. (a) STPa cells could integrate information about the direction of overall displacement during movement (from the dorsal inputs) and information about the form of the stimulus (from the ventral inputs). (b) Sufficient motion information might be available to STPa cells (from the dorsal route alone) to establish sensitivity to the patterns of articulation. (c) Selectivity for body movement could be established by combining inputs from multiple cells (in inferotemporal cortex, IT, or STPa) each selective for the same body form but at slightly different spatial positions. Motion sensitivity could in this case derive from inputs from the ventral route alone using circuitry analogous to that proposed for other systems (Barlow & Levick 1965; Torre & Poggio 1978). This latter mechanism is the least likely, since cells in the STPa and IT have very large receptive fields (Bruce et al., 1981) though changes in sensitivity to stimuli at different positions within the large receptive fields (Gross 1992) could perhaps be used.

Selectivity which could be used in all three processing schemes has already been documented: cells selective for the static form of the head and body are found within the STPa and IT (Bruce et al., 1981; Desimone et al., 1984;

Hasselmo et al., 1989; Perrett et al., 1982, 1984, 1985a, 1991, 1992) as are cells selective for direction of motion but lacking form sensitivity (Gross et al., 1972; Bruce et al., 1981; Perrett et al., 1985b; Hikosaka et al., 1988; Hietanen & Perrett 1993; Oram et al., 1993). Utilization of these cell types could support scheme (a). Inputs to the STP from MT and MST are likely to convey motion information but relatively little form information. Area MT and MST contain increasing numbers of cells selective for the direction of motion independent of local contour motion (Albright 1984; Albright et al., 1984; Duffy & Wurtz 1991a,b; Komatsu & Wurtz 1988a,b; Mikami et al., 1986a,b; Rodman & Albright 1989; Saito et al., 1986, 1989; Snowden et al., 1991, 1992; Tanaka et al., 1986, 1989; Tanaka & Saito 1989; Zeki 1974). These inputs could support scheme (b) or be used in conjunction with inputs from cells selective for static form as in scheme (a). Furthermore, IT, which projects to the STPa (Felleman & Van Essen 1991; Young 1992), contains cells which are selective for static form and perspective view of the head (Tanaka et al., 1991; Hasselmo et al., 1989; Young & Yamane 1992). These cells (and also cells within area STPa selective for form) could be used to support scheme (c) or be combined with inputs from cells selective for motion to support scheme (a). Therefore any of the proposed schemes of processing to detect walking bodies under normal lighting could in principle be implemented by cells in the STPa, using either inputs from cells within area STPa or inputs from cells in areas IT and MT/MST. It is worth noting that under scheme (a), the suggested motion input carries only overall translation information. Therefore form information could not be calculated under biological motion conditions since it is only the motion of the light points relative to one another that can be used to extract form-from-motion (e.g. both left and right profiles moving the left have the same large field motion signals). Similarly under scheme (c) only overall translation information would be present, so again no sensitivity to biological motion stimuli would be seen. Hence STPa cells would be

expected to respond to biological motion stimuli only under the second scheme of processing.

View specificity

The majority of processing of static form information within the STS and IT cortex appears to be conducted in a view specific manner (Bruce et al., 1981; Desimone et al., 1984; Perrett et al., 1985a, 1991, 1992; Hasselmo et al., 1989). For example, individual cells respond to the left profile view of the head but not the right profile or other views. Such sensitivity to perspective view has also been observed in STPa cells conjointly sensitive to body form and motion; some cells respond selectively to the left profile body view walking to the observer's left (Perrett et al., 1985b, 1990a,b). Neuronal sensitivity to the visual patterns of monkey ambulation in specific directions has been observed in other regions of the macaque temporal lobe (e.g. the amygdala, Brothers & King 1992). The view sensitivity seen in STPa cells offers an opportunity to quantify sensitivity to form defined by motion, since the response to movement of one body view can be compared to a different view moving in the same direction. Mirror image body views are identical in size, complexity of articulating elements and angular speed of component movements. Discrimination of responses to different views therefore indicates sophisticated processing of form. Discriminating body view under biological motion conditions has been used in psychophysical tasks to assess quantitatively human perceptual sensitivity to form defined by motion (Cutting et al., 1988; Mather et al., 1992; Perrett et al., 1990a).

Computational approaches for interpretation of biological motion stimuli

The majority of the computational models analysing form from motion stimuli in general (Ullman 1978; Hildreth & Koch 1987) and biological motion in particular (e.g. Rashid 1980; Webb & Aggarwal 1982; Hoffman & Flinchbaugh 1982; Sugie & Kato 1987; Sugihara & Sugie 1984) establish a linkage structure that is consistent for the motion of the moving elements. This means that the

analysis establishes a description that is independent of body view and direction of motion: indeed the analysis is applicable to any articulating entity, and gives no information about direction of movement nor the identity of the object that is moving. While analysis of overall displacement of an object is simple to perform, there would still be the problem of binding the direction of motion with the particular object. Furthermore the models that establish only linkage structure would require a further processing stage for determination of object identity and body view. Other computational approaches use a template of the object's identity (Lee & Chen 1985; Leung & Yang 1987). However, these models also predict invariance with respect to orientation, body view and direction of motion (assumptions are often required about the nature of the motion - walking, running etc - in these models as well). Therefore to explain human observer performance further processing stages would also be required, in particular to extract body view. It would only be during this subsequent stage that effects such as increased response latency seen with inversion of stimuli (Sumi 1984; Dittrich 1993). View sensitivity of cells responsive to body motion defined under biological motion is therefore an important attribute to quantify since it is a property that is not predicted on the basis of most current computational approaches to biological motion.

The primary aim of this study was to determine whether cells in the STPa selectively responsive to the sight of walking bodies under normal lighting conditions were sensitive to biological motion versions of the same walking stimuli. Earlier reports suggested that STPa cells might indeed utilize patterns of articulation (Bruce et al., 1981; Perrett et al., 1990a,b) but no systematic study had been made of the extent to which cell sensitivity to body form and direction of movement was maintained under biological motion conditions.

METHODS

Four subjects were used (Macaca mulatta, 3 male B, D, H wt. 5-8 kg, 1 female J wt. 4 kg from a UK registered breeding colony). The standard training, surgery, recording, behavioural task and histological reconstruction methods were used (see General methods section).

Stimuli

The stimuli were either real 3-D presentations or sequences of frames on a video disc. They included images of the experimenter walking, both forwards (compatible movement) and backwards (incompatible movement) in different directions (towards and away from the monkey and moving to the monkey's left and right). The biological motion stimuli were made using luminescent patches (subtending approximately 0.2 degrees) fixed to the experimenter at the neck, shoulders, elbows, wrists, hips, knees and ankles. Live presentation was performed under black out conditions within the laboratory. Video images were taken both of actors walking under strong diffuse lighting and under blackout conditions to give normal walking stimuli and the equivalent biological motion stimuli. The biological motion stimuli were then contrast thresholded to two luminance levels and finally contrast enhanced to black and white using a Fairlight Computer Video instrument. Both these and the images under natural lighting conditions were stored on video disk.

In addition to the small dot stimuli, stick figure representations were also used. These were generated in an analogous fashion to the biological motion stimuli but short strips of luminescent material were fixed between the articulation points. Gaps of similar size to the `dots' were left at the articulation points. These stick figures have more information than traditional biological motion stimuli

since they give linkage structure but they do *not* have other form information (e.g. appearance of the face).

Control objects moving in the same directions as the walking / translating and biological motion stimuli were used. These were matched for size and like walking had non-rigid motion (e.g. curtains, lab-coats, hinged pieces of wood) and were moved at the same speed (+/- 15%) and direction (+/- 10 degrees) as the walking stimuli. The responses of STPa cells to stimuli of the translating body will be reported elsewhere (Oram et al. in prep). Three biological motion controls were also used. One was luminous dots moving non-rigidly under blackout conditions. Secondly, images of rigid translating dots were stored on video disk. The third was a `jumbled' biological motion figure recorded on video disc. The jumbled figure was made using a computer based system (IRIS 3130, Silicon Graphics). The position of the points of limb articulation were digitized for each of 24 video frames making one step cycle in 4 directions on a treadmill. To create normal motion sequences these were re-animated at 24 frames/sec and each point was allocated an additional translation vector to re-create walking motions with displacement (i.e. walking to the left). For jumbled figures the coordinates were moved in a random direction by a distance that was 30% of the initial head to floor height of the figure. The appropriate motion vector of each of the points was then added as was the translation vector. Therefore the resulting interpretation of the linkage structure was not changed but, when replayed, even though the overall translation and the individual component motions remained consistent with a walking stimulus, the image was no longer recognizable as a human figure. It is important to realize that the light points can still be 'connected' by rigid limb elements and that the component motions were identical to the equivalent biological motion stimulus, but the relative lengths, the relative positions and the relative motions of these elements were no longer humanoid.

The luminance values of the video disk images were 0.2 cd/m^2 for the background, 3.0 cd/m^2 for the dots and stripes (see below) and 4.0 cd/m^2 for the

natural image of a walking person. Under live conditions, the dot and stripe luminance was less than 0.1 cd/m^2, whereas for natural images of a walking person the luminance was 1-4 cd/m^2.

The stimuli were viewed through either a liquid crystal shutter (Screen Print Technology) or a large aperture camera shutter (Compur, 6.5 cm diameter). Both shutters had rise times of < 15 ms. The time at which the shutter became transparent or was fully open was also recorded. Subsequent analysis was linked to this, the true stimulus onset time. Each stimulus was presented 5 or more times in computer controlled pseudo-random order. In addition, a "no stimulus" condition was also used, where only the LED and wall could be seen. This was used to assess background or spontaneous activity (S.A.) levels of the cells. Each trial consisted of a 0.5 s warning tone, followed by a 1 s stimulus presentation period. The inter-trial interval was randomly varied between 0.5 and 5 seconds. Motion of the stimulus was started before the presentation period and continued for a short duration afterwards to ensure a smoothly moving presentation.

The isolated cells were tested using normal lighting and biological motion conditions. Tests were at the least of 2 body views and controls moving in 1 or 2 directions. If, under biological motion conditions, no response to the stick figures was found, then it was assumed that the dot figures would not elicit a response. Cells selectively responsive to body motions other than walking were tested for biological motion sensitivity with luminous patches only under laboratory blackout since appropriate images were not available on video disc. Body motions used in this testing included rotation, crouching and bowing. Cells were also tested for selectivity to the single limb movements present in the preferred stimulus. These tests involved the presentation of the arm or leg flexing and extending in isolation (i.e. rest of the body occluded from sight or visible but stationary). Cells found selective for these stimuli have been reported previously (e.g. leg, arm or hand motion, see Perrett et al. 1985b, 1989a,b, 1990a,b; Mistlin

and Perrett 1990). The cells reported here, however, were unresponsive to individual limb movements.

Data analysis

Analysis of the spike frequency data was performed on-line as a 1-way ANOVA with each condition tested as a factor. The results of this analysis were used to guide subsequent testing. A cell was classified as selective for walking if there was a significant overall effect of conditions and one direction / body view combination was different ($p < 0.05$) from (i) control objects, (ii) a second body view moving in the same direction and (iii) the same body view moving in a second direction. All the cells reported here were not found to be selective for single limb articulation but rather required whole body motion. Cells found to be selective for walking stimuli were then tested with both real walking and biological motion stimuli and subjected to off-line analysis. Off line analysis for all cells took the form of 2-way ANOVA, with the direction of motion as one factor and the stimulus type (natural, biological motion, control) as the second factor. A second 2-way ANOVA was performed with body view as one factor and stimulus type (natural, biological motion) as the second. Significance for all statistical tests was taken at the 0.05 level. Post hoc testing of the ANOVAs was performed using the protected least significant difference (PLSD) test (Snedecor & Cochran, 1980).

RESULTS

Cells Selective to Human Walking

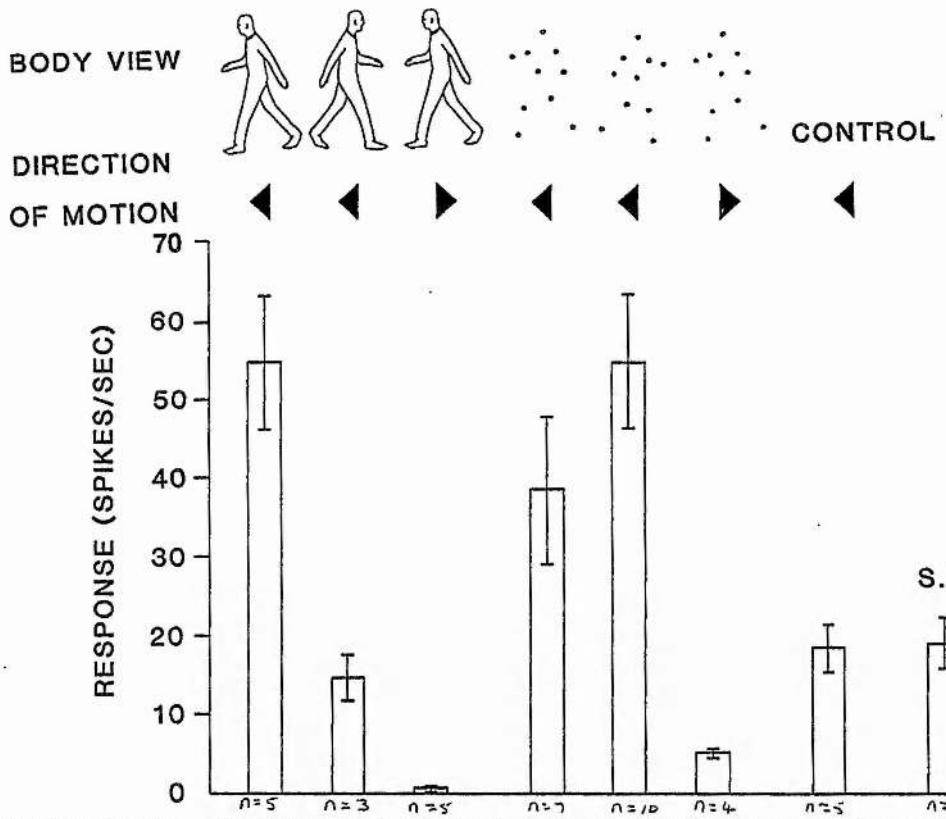


FIGURE 6.1. SELECTIVITY FOR DIRECTION OF BODY MOVEMENT BUT NOT BODY FORM FOR BIOLOGICAL MOTION STIMULI. The mean response +/- S.E.M. of one cell to walking bodies under normal and biological motion conditions. The upper section shows schematic representations of the view and direction of movement of stimuli. Under normal lighting, the cell responded to the sight of the left profile body view walking (compatibly) to the monkey's left. The sight of the right profile view walking (incompatibly) to the left gave no response above the cell's spontaneous activity (S.A.) ($p > 0.05$) and the optimal body view moving to the right produced inhibition relative to S.A. ($p = 0.03$). Under biological motion conditions both left and right body views moving in the cell's preferred direction elicited a response greater than control movement and S.A. ($p < 0.02$ each comparison). Left and right directions of motion were discriminated under biological motion conditions. Overall effect of conditions, $F_{[7,34]} = 14.0$, $p < 0.0005$: Number of trials for each condition, left to right N = 5,3,5,7,10,4,5,3.

From the four subjects, 196 cells were found to be sensitive to walking stimuli out of a total of 6,459 cells screened (see General Methods, chapter 2). Here the response properties of a subset of these 196 cells are described (other cells selective for walking stimuli were subjected to studies of tuning for view, direction and object-part interactions). A total of 72 of the cells found to be selective for the walking stimuli were tested for sensitivity to biological motion (dots and/or the stick figure variation). The selectivity in the responses for human walking could not be attributed to single limb articulation for any of these cells (see methods). Of the 72 cells selective to walking stimuli, 47 (65%) gave no response above spontaneous activity or control response levels when tested with biological motion stimuli. Thus approximately two thirds of the cells selective for walking bodies did not show any responsiveness to stimuli where only motion information was available for defining stimulus form. The lack of response indicates the conjoint selectivity shown by these cells: namely that both form and motion information are required to elicit a response (Oram et al. in prep).

Seven cells (10%) showed a maintained directionality but not view discrimination under biological motion conditions. That is, with moving light displays these cells responded more strongly to both body views moving in the cell's preferred direction than to controls moving in the same direction, spontaneous activity or biological motion in the null (opposite) direction. For instance, in Figure 6.1 it can be seen that the cell does not maintain the view discrimination seen under normal lighting but responds well to biological motion representations of left and right body views moving to the left. Thus sensitivity to body view was not seen. The cell does however maintain direction discrimination. More importantly responses to biological motion stimuli moving left were greater than responses to controls moving in the same direction, indicating partial sensitivity to body form. The responses to the overall direction of the biological motion stimulus did show clear discrimination between movement to the left and

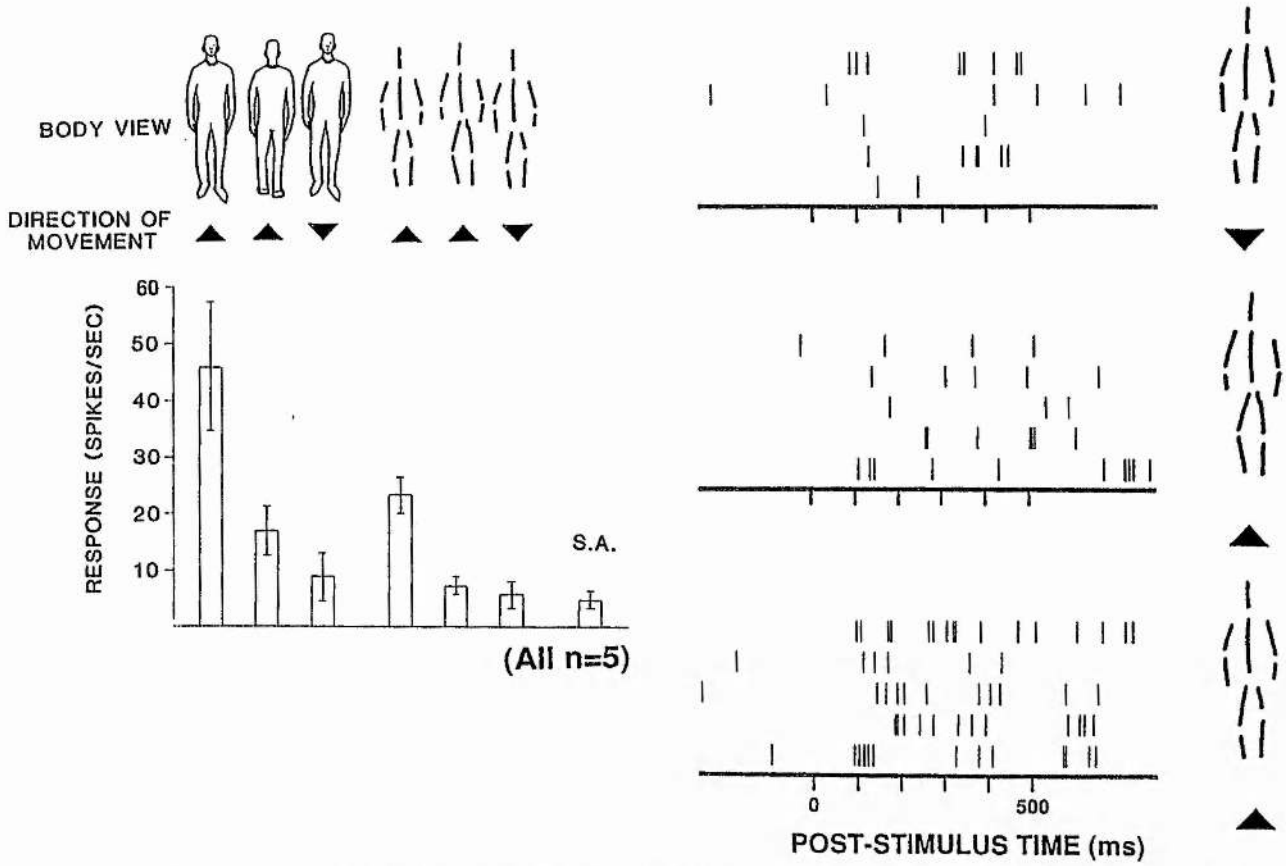


FIGURE 6.2. REDUCED RESPONSE TO STICK FIGURES. LEFT: The mean response +/- S.E.M. of one cell to walking bodies under normal and biological motion (stick figure) conditions. The upper section shows schematic representations of the view and direction of movement of stimuli. Upward arrows indicate movement away from the monkey, downward arrows indicate movement towards the monkey. The view and direction of the stick figures maintain the same order as the normal stimuli. Under normal lighting the cell responded more to the front view walking away from the monkey than to other view/direction combinations, including back view approaching (not shown) ($p < 0.05$ each comparison). The responses to the stick figure walking stimuli gave statistical discrimination between view and direction combinations ($p < 0.05$) that mirrored the responses to the real stimuli but were reduced. S.A. = spontaneous activity. 2-way ANOVA, overall effects: testing conditions (normal vs stick figure) $F_{[1,52]} = 5.53$, $p < 0.025$; view/direction (front/rear, approach/retreat) $F_{[3,52]} = 8.56$, $p < 0.005$; interaction $F_{[3,52]} = 0.673$, $p = 0.57$. Number of trials per condition n = 10 for normal stimuli, n = 5 for stick figures. RIGHT: Rastergram displays of the 5 trials for each of the biological motion stimulus conditions. Each trial is represented by a single row of ticks, each tick indicates one action potential. Post-stimulus time is given at the figure base. Note the latency of approximately 100 ms.

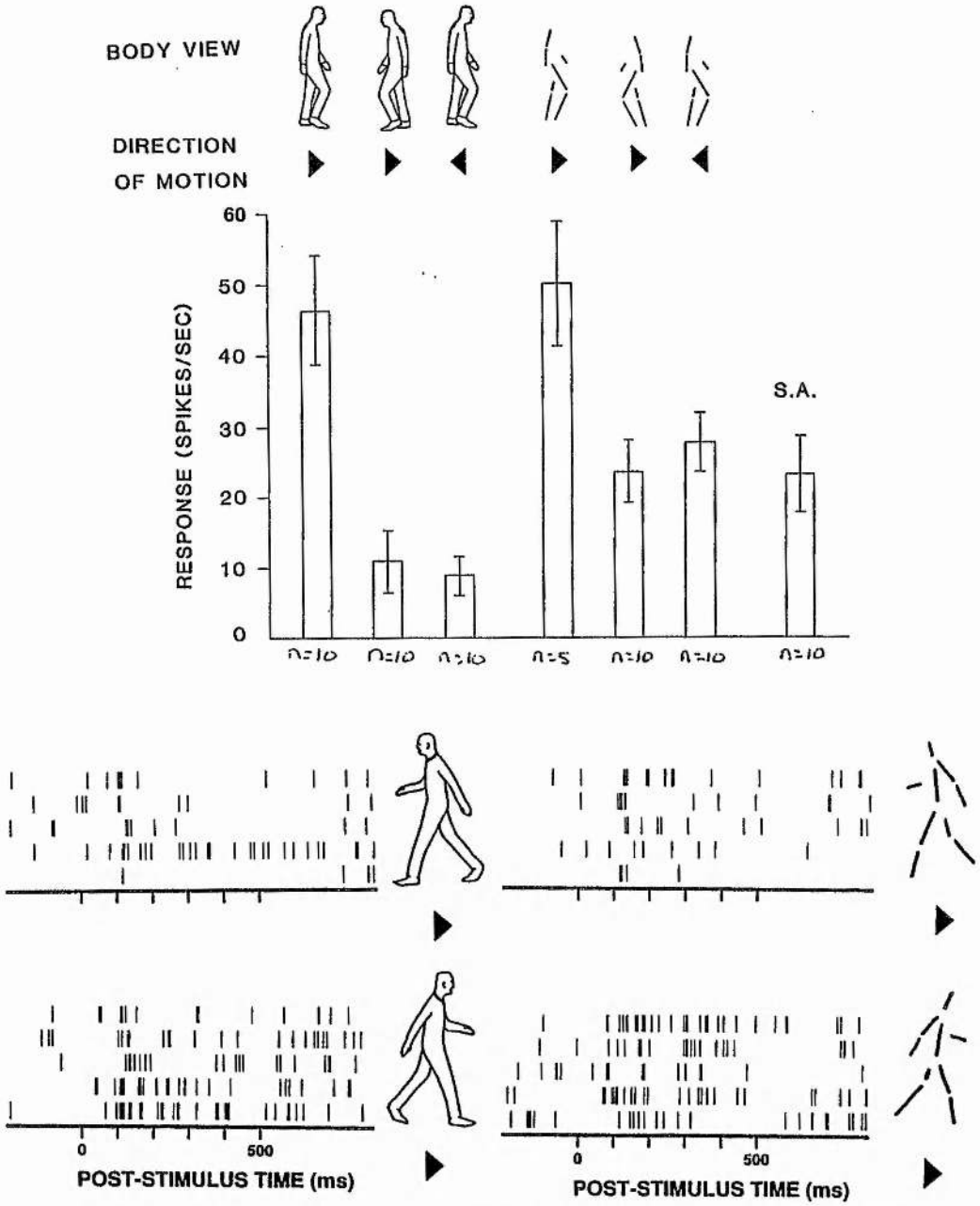


FIGURE 6.3. RESPONSIVENESS TO STICK FIGURE STIMULI. UPPER: The mean response +/- S.E.M. (n=10,10,10,5,10,10,10) of one cell to walking bodies under normal and biological motion (stick figure) conditions. The upper section shows schematic representations of the view and direction of movement of stimuli. The cell responded to the right profile walking to the right, both for normal and stick figure stimuli. Responses to inappropriate views or directions were not significantly different from the spontaneous activity (S.A.) or control stimuli (not shown). 2-way ANOVA showed a main effect for four tested view/direction combinations ($F_{[3,62]} = 11.3$, $p < 0.0005$) but not for lighting condition (normal vs sticks) ($F_{[1,62]} = 1.1$, $p = 0.29$). The interaction was non-significant ($F_{[3,62]} = 2.5$, $p = 0.07$). **LOWER:** Rastergram display showing 5 responses to normal lighting condition stimuli (left) and the responses to the biological motion equivalents (right).

right. No cells were found with the converse selectivity - that is, showing view selectivity but not directional selectivity under moving light displays.

The remaining 18 cells out of the 72 tested (25%) showed full selectivity for biological motion of dot and stick figures. Full selectivity to biological motion is defined here as selectivity for both form (body view) and direction of motion. Ten of these cells were sensitive to moving dot stimuli and 8 to stick figure stimuli. Four cells were tested with both moving dot and stick stimuli. Two of the four gave responses to moving dot stimuli that were statistically indistinguishable from responses to stick figure stimuli, the remaining two cells responding only to stick figure stimuli.

Most of these cells (14/18, 78%) which showed statistical discrimination between directions and body views also showed a reduction in absolute response magnitude relative to the walking stimuli under natural lighting. Cells with this type of response characteristic were found for both moving dot and stick figure stimuli. Figure 6.2 shows an example of this type of response to stick figures. The cell was selective specifically for the front view of the body walking away from the monkey (incompatible movement). With stick figures, a reduced response was found (latency approximately 100 ms) to the preferred stimulus but it was still significantly above spontaneous activity and controls (not shown). Stick figure equivalents of non-effective walking stimuli (e.g. the back view of the body moving away from the monkey) produced significantly smaller responses.

Four cells (6% of the total sample of 72 cells, 22% of the cells responding to biological motion stimuli) responded to the biological motion stimuli in a manner that was very similar to the responses to the real walking stimuli. Figure 6.3 shows the responses of one cell to real and stick walking figures. As can be seen, the cell has a preferred stimulus of compatible walking to the monkey's right. The left profile walking in the preferred direction and the preferred body view walking to the right both produced significantly weaker responses. The stick figure responses also followed this pattern, with no significant differences found

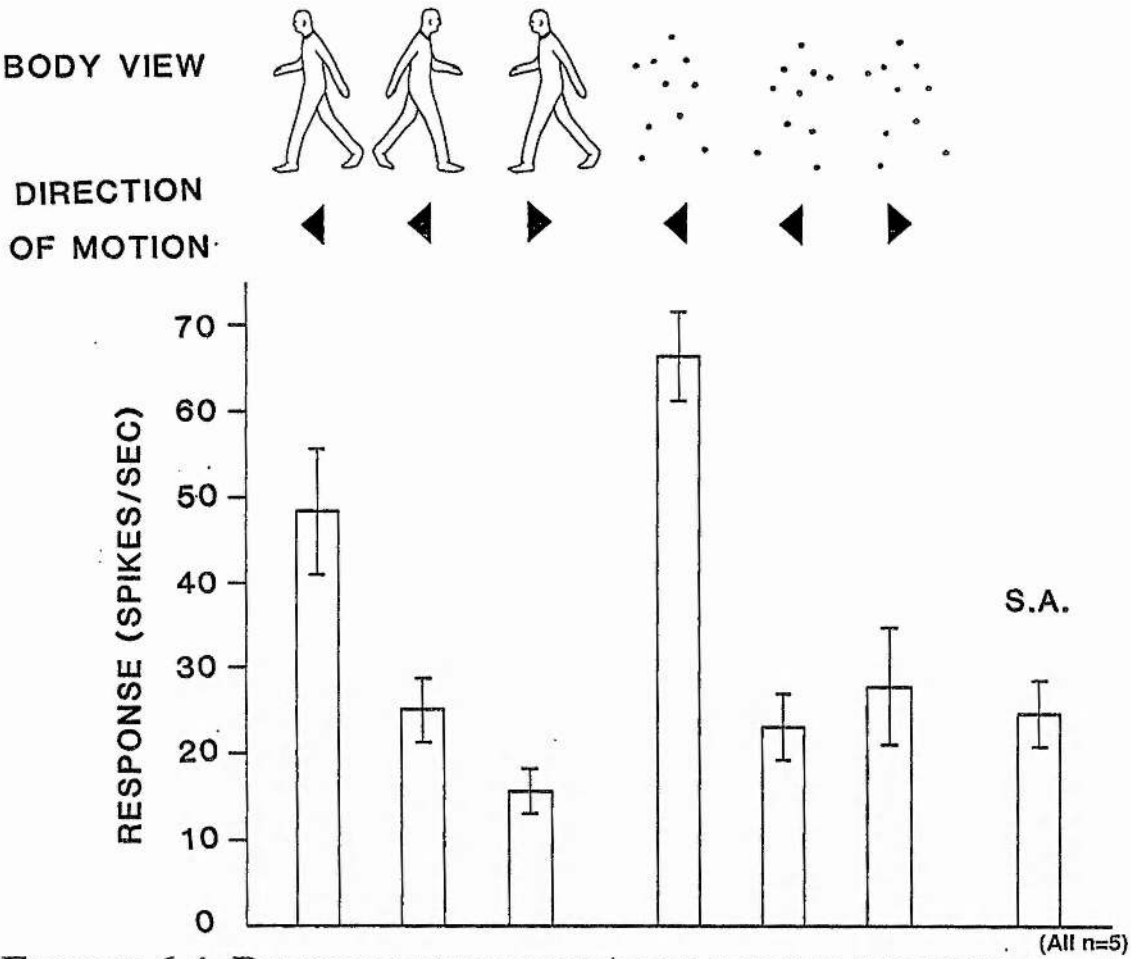


FIGURE 6.4. RESPONSIVENESS TO BIOLOGICAL MOTION. The mean responses +/- S.E.M. of one cell are shown to the stimuli depicted above. The cell's selectivity for compatible walking to the left is maintained with biological motion stimuli. 2-way ANOVA showed a main effect for view/direction combination ($F_{[3,32]} = 14.0$, $p < 0.0005$) but not lighting conditions (natural vs biological motion dots) ($F_{[3,32]} = 3.1$, $p = 0.09$). The interaction was non-significant ($F_{[3,32]} = 1.16$, $p = 0.34$).

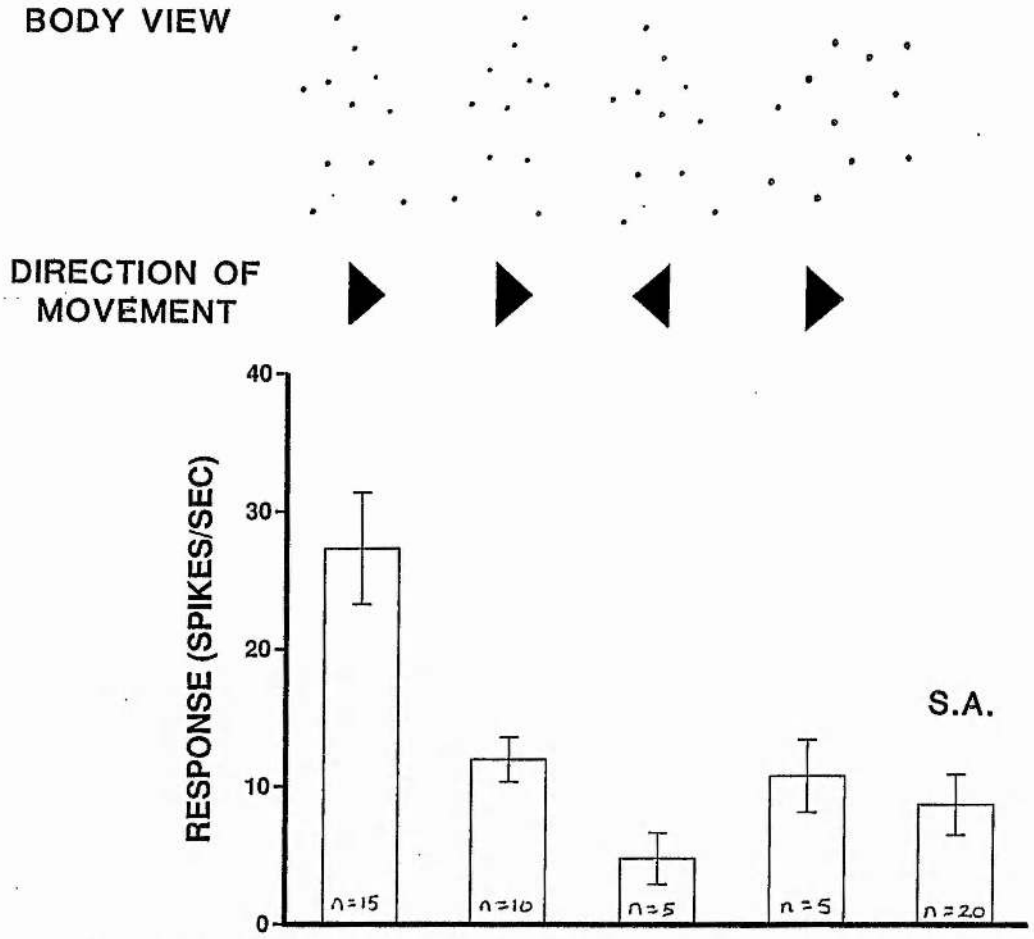


FIGURE 6.5. DISCRIMINATION BETWEEN NORMAL AND JUMBLED ARTICULATION. The upper section depicts the stimuli. Under biological motion (and natural) conditions the cell responded selectively to the left profile view of a body walking backwards. Response to the jumbled biological motion stimulus moving to the right was no different from the response to the moving point light display of the right body view walking to the right ($p > 0.5$) but less than the biological motion representation of the preferred stimulus under natural conditions ($p = 0{:}003$). (1-way ANOVA, overall effect of conditions: $F_{[4,50]} = 8.15$, $p < 0{:}0005$.)

across comparable body view / direction of movement combinations between real and stick figures. As can be seen, the response latency under both biological motion conditions and normal lighting is approximately 100 ms. Figure 6.4 shows the responses of another cell to real and dot figure stimuli. This cell was selective for the left profile view moving to the left. As can be seen this selectivity was maintained at comparable levels when biological motion stimuli were used.

Jumbled Articulation

As an additional investigation of form selectivity, comparison was made of responses to natural and jumbled configurations of the biological motion stimuli. The jumbled figure stimuli (see methods) contain the same rigid linkage structure as the biological motion stimuli, the same overall translation vector and the same component vector of each point: they differed only in the relative positions of the light points. A total of 14 cells were tested with these randomised moving dot displays (jumbled figure). Of these 10 cells proved to be insensitive to biological motion stimuli and the jumbled figure. For the 3 cells where a selective response was seen to the biological motion stimuli (i.e. preferentially responding to one body view and direction combination displayed in biological motion conditions), the response to the jumbled figure moving in the preferred direction was significantly ($p < 0.05$) reduced compared with the preferred view and direction combination. The one cell which was selective for direction but not body view under biological motion conditions also responded to the jumbled figure moving in the preferred direction. Figure 6.5 shows the response of a cell to biological motion stimuli and a jumbled biological motion stimulus. As can be seen, the response differentiates the preferred movement (left profile walking to the right) from compatible movements to the right and the left profile walking left. The response to the biological motion stimuli was reduced compared with the comparable real stimuli (approximately 50%). The jumbled articulation stimulus produced a response that was no greater ($p > 0.5$) than the cell's spontaneous

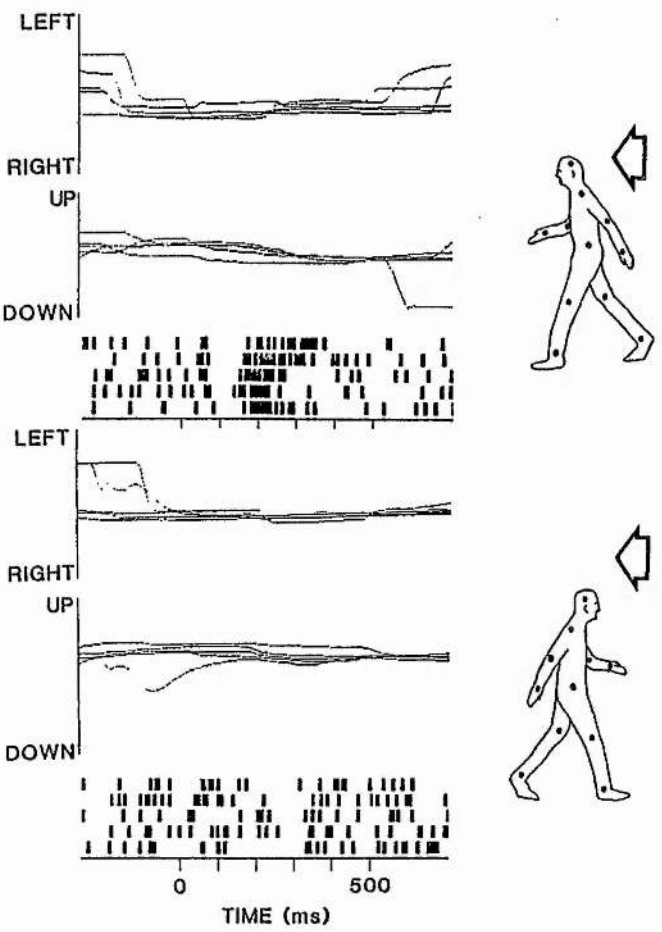


FIGURE 6.6. EYE MOVEMENTS DO NOT ACCOUNT FOR SELECTIVITY OF BIOLOGICAL MOTION. The cell responded to compatible walking to the left. Responses to individual trials of biological motion stimuli are shown in the lower sections, whilst both horizontal and vertical eye movements are shown for each trial in the upper sections. Full scale deflection = 50 degrees for horizontal and vertical traces (eye movements were recorded over a range of positions +/-20 degrees from straight ahead; clipping occured outside this range). The rastergrams show a good response to each of the trials with compatible walking (left profile walking left) but no response to any of the incompatible walking trials (right profile body view walking left). For both stimulus conditions the eye movements were comparable and thus cannot account for the large differences in response magnitudes.

activity. Thus this cell shows statistically reliable discrimination between not only two alternative body view representations but also between the preferred view and direction combination (as a biological motion stimulus) and the jumbled figure equivalent.

Eye Movements

Figure 6.6 shows a typical example of the eye movement recordings and the response to individual trials for one cell. The upper figure shows the responses to an effective stimulus (for this cell biological motion walking compatibly to the monkey's left). As can be seen from the eye movements, just prior to the stimulus onset (time 0) the monkey saccades to the LED and maintains fixation until approximately 250 ms post-stimulus onset. For each of the five presentations, the cell response occurs with a latency of approximately 150 ms. The lower figure shows the eye movements and cell responses to the five trials of the incompatible movement (i.e. walking backwards to the monkey's left). Again, the eye movements show maintained fixation but there is clearly no cell response in any trial. Indeed there is some evidence for inhibition to this type of stimulus. In both the compatible and incompatible stimulus conditions, there is evidence for smooth pursuit eye movement (following the "wrist dot" down) for both stimuli after the initial fixation period. Since the eye movements are comparable for both stimuli, they cannot account for the difference in response magnitudes: the only difference is in the presented stimulus. Furthermore, for the effective stimulus it can be seen that despite small variations in eye position the response onset is tightly time locked (Figure 6.6, upper). A similar lack of relationship between eye movements and response selectivity was obtained for cells from all four recording subjects.

Discrimination Measures

In order to examine the discrimination shown by the responses of the tested cells to the differing stimuli, discrimination measures for both direction (I_d)

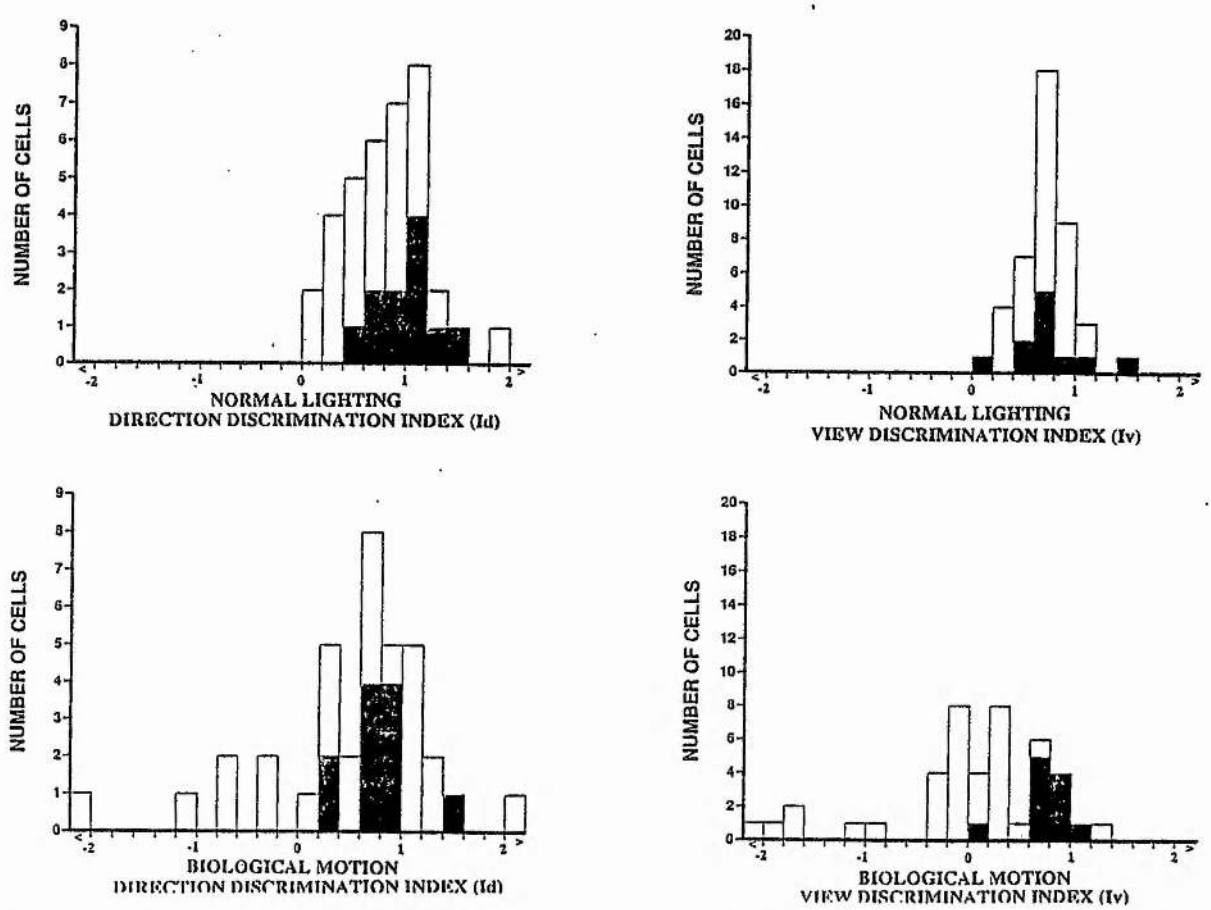


FIGURE 6.7. DISTRIBUTION OF VIEW AND DIRECTION DISCRIMINATION INDICES. LEFT: Distribution of the direction discrimination (I_d) index for all cells selective for body view and direction and tested under natural lighting (upper) and biological motion (lower) conditions. [$I_d = 1 - (R_{oppd}/R_{pref})$, where R_{pref} = response to preferred view and direction - spontaneous activity, R_{oppd} = response to the preferred view moving in the opposite direction - spontaneous activity]. RIGHT: Distribution of the view discrimination measure (I_v) for cells tested under natrual lighting (upper) and biological motion (lower) conditions. [$I_v = 1 - (R_{oppv}/R_{pref})$, where R_{pref} = response to preferred view and direction - spontaneous activity, R_{oppv} = response to opposite view moving in the preferred direction - spontaneous activity.] The black bars indicate cells wich showed statistical discrimination between view and direction under biological motion.

and view (I_v) were calculated. These were calculated as $I_d = 1 - (R_{oppd}/R_{pref})$ and $I_v = 1 - (R_{oppv}/R_{pref})$, where R_{pref} = response to the preferred view and direction combination - spontaneous activity (SA), R_{oppd} = response to the preferred view moving in the opposite direction from the preferred direction - SA and R_{oppv} = the response to the view opposite to the preferred view moving in the preferred direction - SA. The preferred direction and view were first defined under normal lighting and then the magnitudes of the responses R_{pref}, R_{oppv} and R_{oppd} measured and compared under biological motion. The distribution of I_d values is shown in 7a and the distribution of I_v values is shown in 7b. The upper sections show the values obtained under normal lighting conditions; the lower section shows the values calculated for the same cells using biological motion stimuli. The black bars indicate the direction (7a) and view (7b) indices for those cells which showed statistical discrimination for direction and view under biological motion. Note that it would be expected that cells lacking selectivity under biological motion would have a wide range of discrimination index values because the cell responses would fluctuate around spontaneous activity levels.

Comparison of the direction indices for all cells revealed a reduction in I_d for biological motion stimuli compared with I_d for natural stimuli (medians = 0.71 and 0.84 respectively, Wilcoxon test, W=196, N=36, p=0.03). A more marked reduction was found for I_v under biological motion compared to normal lighting (medians = 0.20 and 0.68 respectively, Wilcoxon test, W=89, N=43, $p < 0.0005$). [Values were included only for cells where the responses differed from spontaneous activity by more than 1 spike/sec.] The same comparisons of the indices were also performed for only those cells whose responses under biological motion stimuli showed statistical discrimination for both view and direction. A small drop was observed in I_d (medians = 1.00 (normal lighting) and 0.79 (biological motion), Wilcoxon test, W=9, N=11, p=0.04). Surprisingly, the view discrimination index I_v under the two conditions showed no significant difference

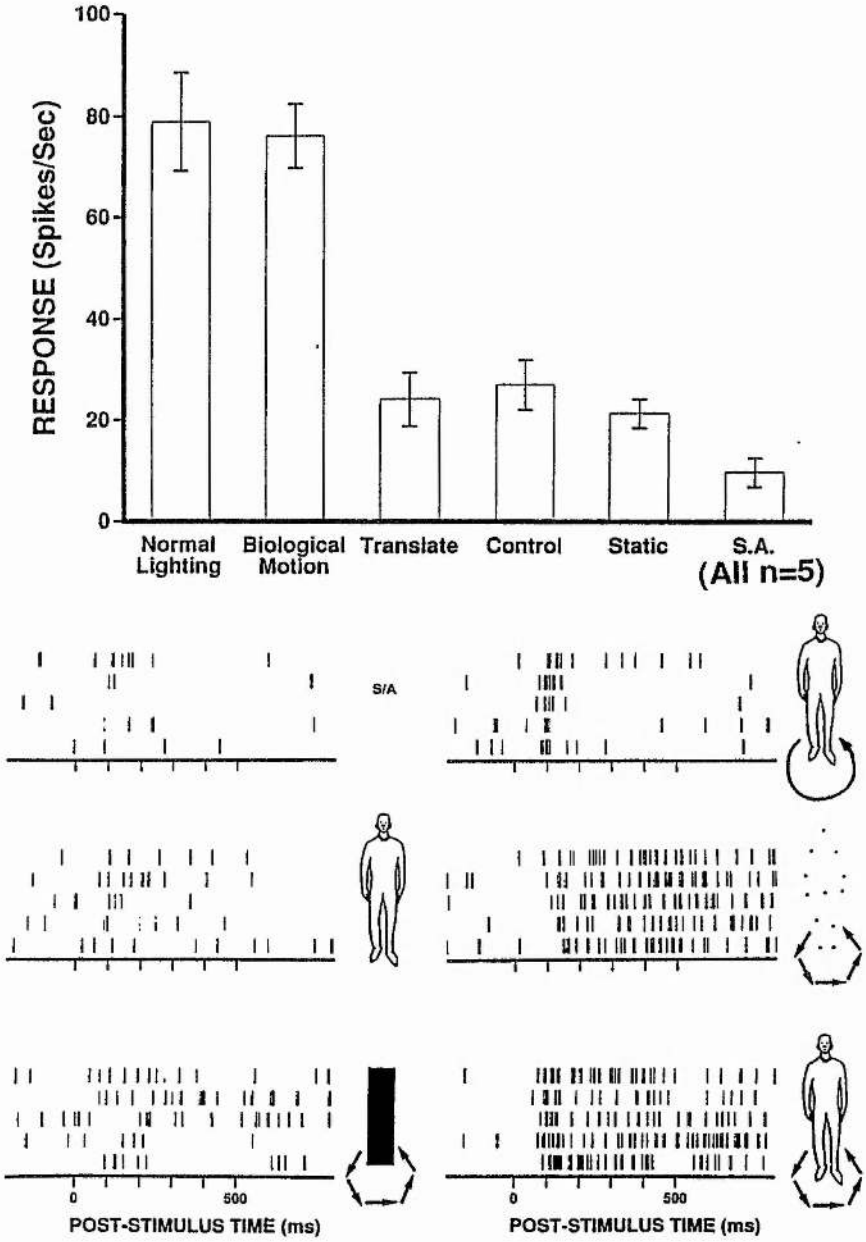


FIGURE 6.8. RESPONSE SELECTIVITY TO BIOLOGICAL MOTION STIMULI FOR ROTATING BODIES. **UPPER:** A single cell that responded to non-rigid whole body rotation under normal and biological motion conditions. Responses were significantly lower to control objects (matched for size) rotating with the same direction and speed, the static body view and a rigid body rotation ($p < 0.0005$ each comparison). There was no significant difference between body rotation under biological motion and normal lighting conditions ($p = 0.73$). Overall effect of conditions: $F_{[5,24]} = 27.1$, $p < 0.0005$. **LOWER:** Rastergram display of the cells responses to the conditions in the upper section. **TOP LEFT:** presentation of the LED alone (S/A or "no stimulus" condition). **MID LEFT:** static body. **BOTTOM LEFT:** control rotating non-rigidly. **TOP RIGHT:** rigid rotation of the body. **MID RIGHT:** biological motion display of articulating body rotation. **BOTTOM RIGHT:** body rotating non-rigidly under normal lighting. The broken arrows indicate non-rigid rotation whereas the smooth arrow indicates rigid body rotation.

(median for natural lighting = 0.69, biological motion stimuli = 0.80, Wilcoxon test, W=27, N=11, $p > 0.5$).

Cells Selective to Whole Body Movements Other than Walking

Sensitivity to biological motion stimuli was found not only for cells selective for ambulation but also for cells during other whole body actions. Ten cells were tested with biological motion stimuli that were responsive to other whole body movements. Five of these were completely unresponsive to biological motion stimuli, whereas the other 5 showed either reduced or comparable response magnitudes and response patterns to those obtained under normal lighting conditions. Figure 6.8 shows an example of a cell that was selective for the sight of the whole body rotating. Whilst the direction of rotation did not matter (not shown), controls of a comparable size rotating at similar speeds with component articulation did not elicit a response. Biological motion stimuli produced responses comparable with the live rotations. This cell was of interest because, when tested with the rotation without articulation (achieved by having the experimenter standing rigid on a rotating platform under normal lighting), the cell showed only a very weak response. This implies that, for this cell, the sight of limb articulation during whole body rotation was necessary. It also was apparent that biological motion conditions were sufficient to elicit a strong response. The rastergram display (Figure 6.8b) suggests that the response consists of two components. The first of these is transient (lasting 10-40ms) and can be elicited by the sight of rigid body rotation. However the following sustained response can only be elicited by non-rigid rotation of the body.

A second example of a cell selective to whole body movement downwards defined by the pattern of articulation which was also sensitive to biological motion stimuli is shown in Figure 6.9. Translation of a non-articulating body (a life sized 2-D model) down produced a non-significant response that was comparable to control movement down. Thus for this cell the articulation was a

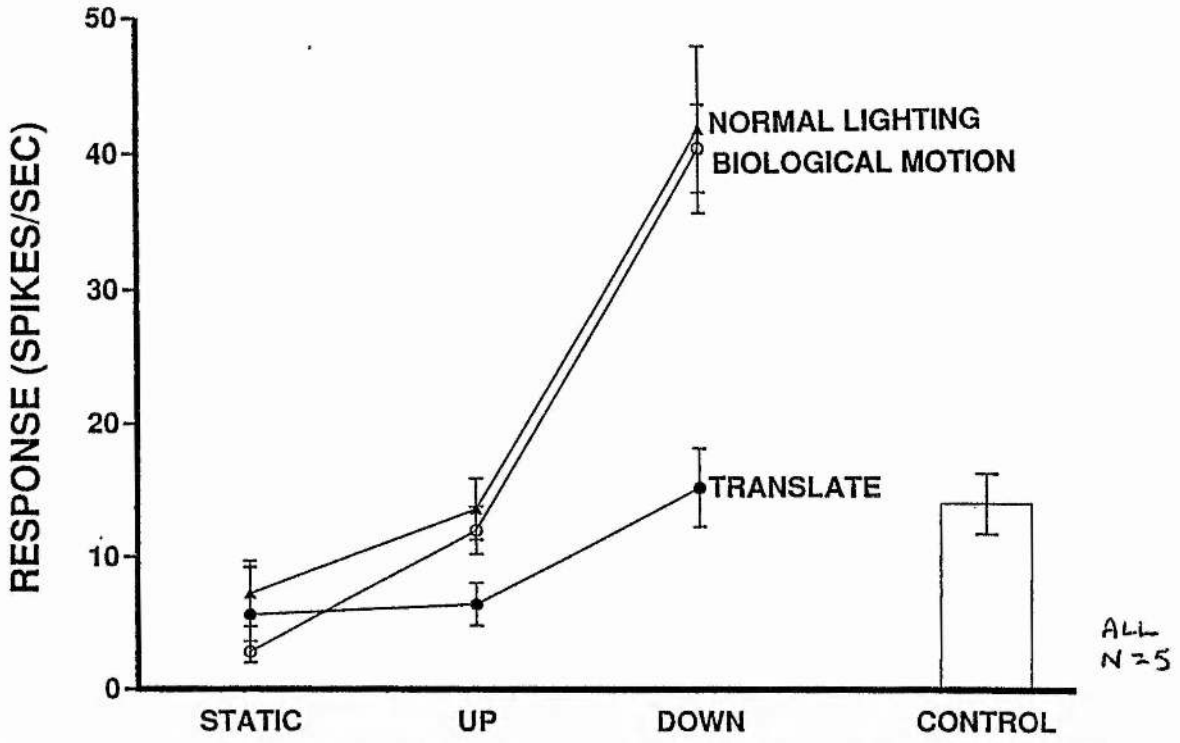


FIGURE 6.9. ARTICULATION AS A NECESSARY AND SUFFICIENT CONDITION FOR RESPONSE TO BODY MOTION. The mean responses (+/- S.E.M., n=5) are shown to the stimuli represented in the lower section. Crouching down, either as a biological motion stimulus or under normal lighting produced a response significantly greater than reverse direction of body motion (standing up), static body views and control movement downwards ($p < 0.05$ each comparison). [2-way ANOVA main effect of motion type (down/up/static) $F_{[2,46]} = 45.0$, $p < 0.0005$; main effect of stimulus type (translate/natural/biological motion) $F_{[2,46]} = 7.89$, $p = 0.001$; Interaction $F_{[4,46]} = 2.66$, $p = 0.045$] Downwards translation of the head and body (without articulation) did not produce a response different from control motion downwards ($t_{[8]} = 0.32$, $p > 0.75$).

necessary component of the stimulus. The cell's response to the biological motion version of the body moving down (with articulation) was similar to the response to the same stimulus under normal lighting. Thus for this cell it was shown that the relative movements between the points of articulation produced by crouching down were necessary and sufficient to produce the maximal cell response.

Location of Cells

Figure 6.10 shows the histological reconstruction of the positions of the tested cells in the central region of the recording area in one subject (monkey J). The left columns show the locations of all cells tested for sensitivity to biological motion stimuli. The right columns show only those cells which responded selectively to biological motion (either dots or stripes or both). As can be seen in the right hemisphere, cells selectively responding to biological motion stimuli were found in both the upper bank and the fundus of the sulcus. In the left hemisphere, the cells tested for sensitivity to biological motion were only located in the upper bank of STPa (areas TPO and PGa of Seltzer & Pandya 1978). There was no reason to assume that the distribution of sensitivity to biological motion stimuli showed any hemispheric differences. Reconstruction of the other subjects indicated that cells sensitive to biological motion were located in the same regions.

DISCUSSION

Summary of the results

One third of the cells (25/72) selective for the form and motion of walking bodies that were tested showed partial or full sensitivity to biological motion (either dot or stripe) stimuli. Of the cells that did respond, one third (7/25) showed only partial sensitivity to form from motion, in so far as the responses under

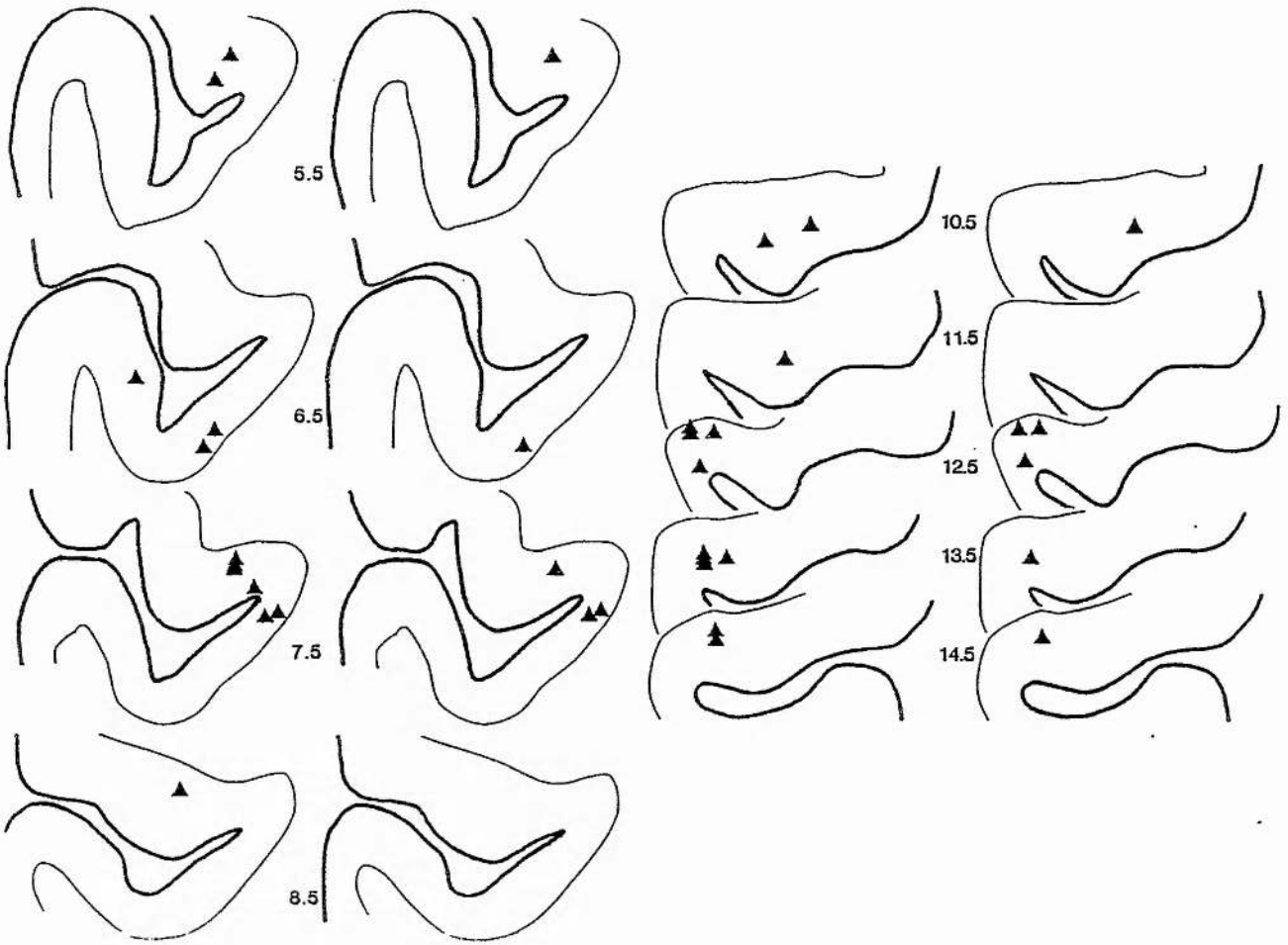


FIGURE 6.10. HISTOLOGICAL RECONSTRUCTION OF RECORDED CELLS. LEFT: Series of 4 sections (at 1mm intervals) of the superior temporal sulcus (STS) from the right hemisphere of one monkey (J). RIGHT: Series of 5 sections showing only the upper bank of the STS from the left hemisphere of the same monkey. The thick line indicates the surface of the brain. the thin line marks the boundary between grey and white matter. The left columns of a and b show the location of the cells tested for sensitivity to biological motion. The column on the right shows the location of those cells which showed statistically significant discrimination for the form of biological motion stimuli. The figures between the two columns indicate the distance in millimetres anterior to the interaural plane.

biological motion conditions maintained direction sensitivity but failed to discriminate body view. For these cells a limited capacity to process 'body form' was indicated by the observation that often responses to all views of the walking body depicted in biological motion were greater than moving dot or stripe control stimuli.

The majority (18/25) of cells responding to biological motion stimuli showed statistical discrimination for both direction of movement and body view. Such cells typically showed reduced responses to the impoverished stimuli compared to full light conditions. Response reduction is perhaps not surprising given the loss of contour information in the biological motion stimuli. Some (4) cells showed response magnitudes and selectivity to biological motion stimuli that were statistically indistinguishable from those to the `real' stimuli under normal lighting. All 18 cells were thus sensitive to detailed form information (body view) from the pattern of articulating motion present in biological motion stimuli. The cell responses thus provide direct evidence for neural mechanisms computing body form from non-rigid motion.

In relation to the possible schemes of processing described in the introduction, the data here presents evidence in favour of the scheme (b) whereby form is calculated from motion inputs alone (in particular see Figures 6.8 and 6.9). Although all three mechanisms outlined in the introduction could contribute redundant information to the ultimate perception of body motion, under biological motion conditions only mechanism (b) would contribute to the perceptual ability to differentiate body form (schemes (a) and (c) would predict only responses equivalent to non-rigid controls moving in the preferred direction). As noted in the introduction, area STPa receives inputs from both areas MST and FST (Boussaoud et al. 1990). However, the number of cells (6%) showing comparable responses to natural lighting and biological motion stimuli is small: the majority of cells responding to the impoverished biological motion stimuli showed reduced response magnitudes.

An associative model for biological motion sensitivity

The observation that most cells show reduced responses to biological motion stimuli, and that nearly all cells responded to translation of the appropriate body form in the preferred direction (Perrett et al. 1990; Unpublished observations) leads to the following tentative model for the derivation of cellular selectivity to biological motion stimuli. This speculate model relies on a simple associative mechanism that might explain the emergence of sensitivity of some STPa cells to biological motion stimuli. While in this chapter the focus has been on responses to biological motion stimuli, most cells in area STPa sensitive to walking bodies also show sensitivity to translation (in the appropriate direction) of a non-articulating image of the body (in the appropriate view) (Perrett et al. 1990). This sensitivity to translation is also found for many of the cells sensitive to biological motion. This suggests that cells in the STPa sensitive to walking receive form information and motion information separately. Further, the sensitivity of some biological motion sensitive cells to translating stimuli also indicates that both local and global motion inputs may influence cell activity.

Cells in STPa could receive inputs about body form (from the ventral stream) and various local motion inputs about limb articulations (from the dorsal stream). When the body locomotes, the sight of *one* body view translating in *one* direction would be associated with *one* set of local motions. This triple association would be the basis for learning the collection of articulations which typify one body view moving in one direction, with one type of action (walking).

The responses of cells to the static form of the body cluster around particular 'characteristic' views (front, back and profiles, Perrett et al. 1991). Further the direction selective neurons in area STPa cluster around 'characteristic' directions (left, right, towards and away, Oram et al. 1993; see also chapter 5). If sensitivity to biological motion is learned associatively then it should occur with the same type of view and direction selectivity. This is exactly what we found.

The data from our studies fit several predictions from the associative learning scheme proposed here: (1) sensitivity to biological motion should show form and direction selectivity, (2) since form inputs are required as a basis for associative learning, responses to biological motion stimuli are likely to be reduced compared with normal lighting stimuli, (3) sensitivity to biological motion should not be present for cells responsive to the form of static bodies (the current data supports this - unpublished observations), (4) cells showing selectivity for walking movements under biological motion conditions should also show selectivity for translation (in the appropriate direction) of the rigid body form (in the appropriate views). The physiological evidence is thus consistent with a scheme of processing in which the local motion of limbs is associatively learned with paired form and global motion information.

In summary, it is suggested that scheme (a) of the introduction is prevalent, with cell selectivity to form and motion (body walking) results from integration of form inputs and motion inputs from separate dorsal and ventral sources. For a minority of cells it is suggested that the motion inputs also include local field (possibly from MSTl) as well as wide field inputs (from MSTd). It is possible that it is the learned association of local field inputs (potentially coding relative motion of light points) with the overall translation (wide field motion) and form inputs that gives rise to sensitivity to biological motion stimuli. While this tentative model would explain many of the results reported here (Oram & Perrett 1993), the cell selectivity to biological motion stimuli reflects the ability of cells in the macaque STPa (after "learning") to compute form from the motion inputs alone (scheme b of the introduction), and does not rely on the presence of form inputs.

Sensitivity to global motion patterns

Populations of cells in the anterior sections of the temporal lobe have been found sensitive to the movement of individual limbs (Perrett et al., 1985b,

1989a,b, 1990b; Hasselmo et al., 1989). The cells reported here, however, responded only to whole body motion and not single limb articulation. It is unlikely therefore that the sensitivity observed to biological motion stimuli can be accounted for in terms of isolated patterns of local relative motion. The global nature of the motion analysis was indicated by the discrimination of body view for whole body movements in the same direction and by the observations that cells were (1) unresponsive to control patterns of dots moving non-rigidly and (2) could respond differentially to jumbled and normal biological motion stimuli. These observations indicate the complexity of the analysis being performed, since all connected pair-wise relative motions of individual limbs remain in the jumbled and opposite view stimuli, yet cells did not respond. An analogous situation exists with some cells selective for static views of the head. These cells discriminate between different views with the same facial features (e.g. left and right profile) and they also respond less to the presentation of a jumbled face even when all the facial features are present (Perrett et al., 1982, 1988, 1991, 1992).

Motion processing in the ventral visual areas

The stream of visual processing running ventrally into the temporal cortex is commonly thought to be associated with the encoding of object form. The specification of an object's form is usually thought to involve an analysis of static visual information. Indeed lesion studies have indicated that temporal cortex is needed for the learning and memory of static patterns (Dean 1976; Ungerleider & Mishkin 1982). Processing of static form in this region is also indicated by the finding of single cells which exhibit a high degree of selectivity for static objects (see introduction).

The results presented here show that neural sensitivity to form and motion does not depend solely on form visible at any particular instant but can be generated from motion information alone (scheme b of the Introduction). The computation of form from motion may well involve or depend on processing

conducted in the dorsal stream of processing. Certainly lesions to the dorsal system (MT/MST) can produce impairment in the extraction of shape from motion (Andersen & Siegel 1989; Siegel & Andersen 1986). The properties studied here could well depend upon the projections from the motion processing areas (MT/MST/FST) into the cortex of the STS (Felleman & van Essen 1991).

Lesions of the inferior temporal cortex in monkey impair the ability to learn shape discrimination where shape is defined by the relative translation of random dot patterns (Britten et al., 1992). This finding suggests the processing of movement information within the ventral stream, and indeed recent studies in V2 and IT have shown that these areas are sensitive to motion defined contours and simple shapes (Peterhans & von der Heydt 1993; Sary et al. 1993). These new findings therefore suggest that selectivity to form from motion might be achieved through the processing of contours defined by motion in the ventral stream. However, if this were the case, then cells selective for biological motion stimuli should also respond well to static images. All the cells contributing to the summary given here responded only to moving stimuli. Furthermore, testing of cells selective for static images of the head and body has not revealed selectivity to biological motion defined form (unpublished observations). It therefore seems unlikely that cell selectivity to biological motion stimuli is achieved through processing of motion defined contours along the ventral stream.

Relation to neuropsychological studies

It is becoming increasingly apparent from neuropsychological studies that recognition using static and dynamic visual cues is dissociable. Impairments in the ability to recognize facial expression from static photographs (Ekman & Friesen 1976) do not necessarily parallel recognition impairments for expression displayed in biological motion format with light dots attached to the face (Bassili 1979; Humphreys et al., 1993; for discussion see also Campbell et al., 1992).

Neuropsychological studies also indicate that human brain mechanisms involved in the processing of complex motion (such as body form defined by biological motion and the form of cylinders defined by rigid rotation) can be dissociated from mechanisms involved in processing of direction and velocity (Vaina et al., 1990). More dorsal lesions are associated with a loss of simple motion processing, whereas lesions more anterior and ventral are associated with disruption of form from motion. A further example of this dissociation is provided by Patient LM (Zihl et al., 1983) who has been described as 'motion blind' following lesions to dorsal visual areas. LM cannot track movements at velocities greater than 8 degrees per second; fast moving objects appear to her as a series of static images. Despite this dramatic motion processing deficit LM retains some capacity to recognize body form defined by biological motion stimuli (McLeod, Zihl, Perrett and Benson, unpublished studies, 1990).

Relationship to computational models

Ullman's algorithm for extracting form from motion could apply to biological motion stimuli except that it requires 4 visible non-coplanar points on each rigid element (Ullman 1979). The earliest computational model to calculate an object's linkage structure from biological motion displays (Rashid 1980) used the correlation of position and velocity of dots in successive video frames to postulate the rigid connecting links between the dots. This simple procedure produced reasonable solutions for simple stimuli (an idealized walking man). For complex stimuli (e.g. 2 men walking around one another) the procedure was slow and inaccurate. More recent computational approaches (Hoffman & Flinchbaugh 1982; Bennett & Hoffman 1985; Sugie & Kato, 1987; Webb & Aggarwal 1982) make use of natural constraints which are likely to exist in the stimuli. For example Webb & Aggarwal (1982) assume the axis of rotation of each locally rigid element remains fixed during the rotation. The resolved trajectory for one rod element (e.g. the torso) can be used as a frame of reference for defining the

trajectory of the next linked rod element (upper arms and leg sections). Such approaches can resolve the correct linkage in biological stimuli extremely efficiently, indeed performance can reach the theoretical limit of 3 successive frames providing no assumptions are broken (e.g. Sugie & Kato 1987). In summary, the assumptions made in the computational approaches outlined above are mostly about the types of motion allowed between the elements. There is a spectrum of models, each making different assumptions. They range from template matching of 3-D structure (Lee & Chen 1985; Leung & Yang 1987) to establishing the projected 2-D image linkage (Rashid 1980). Although many models can theoretically calculate 3-D structure efficiently (within 3 snap shots or frames of motion) they are not robust but rather sensitive to failure in the assumptions (for example see Webb & Aggarwal 1982). Further, correspondence of light points between frames and velocity information is often assumed as part of the input data, information that is rarely available from just 3 frames (Rashid 1980).

It is relevant to consider the data from psychophysical studies which indicate that naive human observers may perform less efficiently than the recent computational models. Naive observers can correctly identify a biological motion stimulus with exposure durations of between 0.1 and 0.2 seconds (4-8 frames, Johansson 1976; Lappin et al., 1980). With computer animated biological motion displays subjects can discriminate normal walking figures from jumbled figures where the position of limb marker points has been moved randomly a distance 30% of the head to ankle height (Perrett et al., 1990a). Observer performance on such discrimination tasks is profoundly affected by the presence and type movement of background masking dots (Cutting et al., 1988; Perrett et al., 1990a; Proffitt et al., 1984), unlike the computational models which should have no problem with masking dots. Naive subjects perform the normal/jumble discrimination task initially rather poorly and often require more than 8 frames to perceive the figures. Minimal practice (30 trials) substantially improves

performance. Even in the presence of background masking dots, which remove residual static form cues, experienced subjects can perform above chance with 2 to 3 frames exposure.

We learn from these perceptual studies in humans that purely dynamic cues can be used to retrieve structure extremely quickly. Considering STPa cell response latencies similar conclusions can be reached. Although detailed studies of the response time course have yet to be made, it is apparent that cell responses to biological motion stimuli can occur within 100-150 ms after stimulus onset (Figures 6.2, 6.3, 6.6, 6.8).

Computational models derived so far for interpretation of biological motion have two properties that make them inadequate for accounting for psychophysical and single cell data. The first property is that the models are general purpose. The perceptual system, however, appears to employ specific mechanisms rather than a general purpose analysis. Sumi (1984) and Dittrich (1993) found that normally oriented biological motion stimuli were more accurately perceived than inverted stimuli. If the visual system employs a general purpose analysis then perception should be equally successful in identifying inverted or upright figures. The physiological data indicate a much higher degree of specialization. Even for the normally experienced upright orientation, different cell populations are employed for the analysis of different types of body motion (walking, crouching, rotating). At a more detailed level, for each type of movement (e.g. walking) sub-populations of cells are involved in the analysis of specific directions of movement and specific body views (e.g. left profile view walking left). In all, 8 sub-populations would be needed to cover bipedal walking along the horizontal plane (two types of motion, forward and backward in four directions left/right and towards/away). Thus while the majority of computational models apply equally to all perspective views, the brain systems involved in computing biological motion appear to employ view and direction specific neural mechanisms.

The second important difference between neural mechanisms studied here and the computational approaches is that many of the models achieve a less complete description of the visual input compared to natural recognition systems. Many computational schemes only retrieve the linkage structure (which element connects with which) whereas the systems studied here are capable of providing additionally information about the nature of the linked stimuli. The cell responses can for example provide evidence that the stimulus is a body (as opposed to other objects or a jumbled body), that it is walking (not rotating or crouching) and more specifically that it is seen from left profile view and is walking to the left.

The improvement of human perceptual performance with practice indicates that the processing of biological motion stimuli may in some way involve 'top-down' influences where expectations for the form of the moving object are compared against visual input. The appropriate computational model for processing would appear to be one in which input data are checked against specific models stored in memory and the results of the matching used to guide subsequent predications (see Lee & Chen, 1985; Leung & Yang, 1987). A role for top-down influences has also been suggested for object recognition (Lowe 1987; Seibert & Waxman 1991, 1992a,b). It remains to be determined what role experience has in shaping STPa cell responses to biological motion stimuli (see above).

Hybrid computational models might be more appropriate for describing the cellular responses to biological motion stimuli. Such models could perhaps first search for potential links in the articulating array and then check these against specific stored representations of the static or articulating bodies (see Lee & Chen, 1985; Leung & Yang, 1987). The stored representations could be object centred (Marr 1982; Marr & Vaina 1982; Marr & Nishihara 1978; Lowe 1987) or, more in agreement with the physiological data, view-point dependent (Koenderink & van Doorn 1979; Seibert & Waxman 1991; Goddard 1992).

CHAPTER 7

INTEGRATION OF FORM AND MOTION IN THE ANTERIOR SUPERIOR POLYSENSORY AREA (STPa) OF THE MACAQUE MONKEY I. EVIDENCE FOR THE BINDING OF VISUAL INFORMATION AND CLASSIFICATION OF RESPONSE TYPES

(Oram & Perrett, 1995a, *J. Neurophysiol.* in submission)

INTRODUCTION

Our every-day experience of visual stimuli is one of a coherent world in which there is no ambiguity concerning which objects in a scene are moving and which are static. Perceptually, therefore, we have no difficulty in binding the visual attributes of form and motion into a coherent whole. This is surprising given the anatomical, physiological and neuropsychological evidence, from both monkey and man, suggesting that visual processing is performed by different functional pathways (see Figure 7.1). The integration of two separate aspects of information about a single object has been referred to as the *binding problem* (von der Marlsburg 1988; Treisman and Gelade 1980; Treisman and Sato 1990). The computational processes hypothesized to solve the binding problem have limited success but show little of the mechanism used (e.g. von der Marlsburg and Schneider 1986; von der Marlsberg 1988; Singer, 1990; Engel et al. 1992a,b; Treisman and Gelade 1980; Yamaguchi and Shimizu 1994). The marked discrepancy in primates between the functional separation of form processing and motion processing and perceptual experience has received little neurophysiological investigation, yet bears directly on the computational

processes involved in solving the binding problem. We briefly review the evidence for this functional division below and argue that the anterior part of the superior temporal polysensory area (STPa) is one of the few cortical areas where these two types of visual information converge (see also Young 1992).

Evidence from anatomical studies of the monkey has suggested for more than a decade that there are two visual pathways in the cortex of primates (Ungerleider & Mishkin 1982; De Yoe & Van Essen 1988; Goodale & Milner 1992). As further studies have been performed the known number of visually responsive cortical areas and the complexity of the connections between them has been steadily increased. At present there are at least 32 known visual areas (Felleman & Van Essen 1991), although this number is likely to continue to rise. The suggestion that there are two pathways (Newcombe and Russell 1969; Ungerleider & Mishkin 1982; De Yoe & Van Essen 1988; Felleman & Van Essen 1991; Merrigan and Maunsell 1993) has been strengthened from an examination of the density of connections between the visually responsive cortical areas (Young 1992). Figure 7.1 shows schematically this division of visual processing and indicates some of the main cortical areas associated with these two pathways.

Physiological evidence for two pathways

Recordings of single units in monkey cortex indicate a separation of visual processing into a form processing pathway and a motion processing pathway. Recordings from V1 suggest that cells within the cytochrome oxidase rich areas (blobs) code wavelength (colour), whereas orientation and motion are predominantly coded by cells in the cytochrome oxidase sparse interblob regions. In area V2 this separation appears to be magnified, with the cytochrome oxidase rich thick stripes containing cells selective for motion, the thinner cytochrome oxidase rich stripes containing wavelength selective units and the interstripe or pale stripes containing cells showing orientation selectivity (Hubel and Wiesel 1968; Michael 1981; De Valois et al. 1982a,b; Blasdel et al 1992a,b; Hubel and

Livingstone 1987, 1990; Shipp and Zeki 1985; Peterhans and von der Heydt 1993). Recently evidence from recordings in V2 of alert monkeys suggests, however, that this division might not be as clear as seen with anaesthetized preparations and that V2 may have little role in the processing of motion direction *per se* (Peterhans and von der Heydt 1993). That V1 and particularly V2 are involved in processing of form is suggested by the finding that some cells in these areas respond to "illusory" or "subjective" contours (Peterhans et al. 1986; von der Heydt et al. 1992; Peterhans and von der Heydt 1989a,b; von der Heydt and Peterhans 1989; Grosof et al. 1993).

We consider first the processing of visual motion information. The motion sensitive cells of V1 send a strong projection to the middle temporal (MT) area (also known as V5, Dubner and Zeki 1971; Zeki 1974). Macaque area MT contains a very large number of motion selective cells (Zeki 1974; Maunsell and Van Essen 1983; Albright 1984; Albright et al. 1984; Movshon et al. 1985) arranged in columns, each column containing cells showing similar directional selectivity (Albright et al. 1984; Saito et al. 1989). The role of area MT in perception of motion direction of motion in macaques has been elegantly shown by micro-stimulation of small number of cells (Newsome et al. 1990; Salzman et al. 1990, 1992; Murasugi et al. 1993). MT neurons respond to motion direction defined by a number of visual cues (e.g. luminance, texture or colour: Albright 1992; Saito et al. 1989). Importantly, in area MT some 30% of cells show directional selectivity that runs in the same direction as the principle axis of bars moved across the receptive field whereas directionally selective cells in V1/V2 show maximal directional selectivity perpendicular to the principle axis of the stimulating bar (Albright 1984). Therefore area MT appears to be the first cortical area in the motion processing or dorsal stream that shows some "form" selectivity in a way similar to the end-stopped neurons of V1 and V2. However, this form selectivity is slight: although cells in MT may generate signals that can be used to establish direction of motion of an arbitrary object (Albright et al. 1984; Snowden

et al. 1991, 1992) the responses do not themselves indicate the identity of the object. A motion processing area in human cortex thought to correspond to MT has been found using positron emission topography (PET) and dipole analysis from event related potential (ERP) studies (Zeki et al. 1991, 1993; Probst et al. 1993; Watson et al. 1993).

The dorsal pathway continues to the dorsal and lateral portions of the medial superior temporal area (MSTd and MSTl respectively). Neurons in MST (MSTd in particular) prefer motion over a wide field (Tanaka et al. 1986, 1989; Komatsu and Wurtz 1988a,b; Tanaka and Saito 1989; Duffy and Wurtz 1991a,b). Cells in MSTl show sensitivity to relative motion of static objects covering a moving background such that it is the relative motion energy of the object's contours that drives the cells' responses (Sugita and Tanaka 1991). Most cells in MST show no response to self-induced retinal motion resulting from eye movements (Duffy and Wurtz 1991b; Erickson and Thier 1991). Sensitivity to disparity has been shown in MST neurons (Roy and Wurtz 1990; Roy et al. 1992). It has been suggested that this area has a role in maintaining visual stability during self motion (Tanaka & Saito 1989; see also Duffy & Wurtz 1991) and in the control of smooth pursuit eye movements (Komatsu & Wurtz 1988a,b, 1989). It should be noted that the large receptive field properties suggest that area MST has little role in signalling the form of a moving object.

Area MST projects to the parietal cortex (ventral and lateral intra-parietal areas and from these areas to area 7a). Cells of the parietal lobe show directional selectivity for visual motion (Sakata 1991; Colby et al. 1993) and seem to be related to arm/hand movement co-ordination (Blum 1985, 1989; Mackay et al. 1992). Neurons of the parietal lobe may also encode a representation of perceptual space (for review see Sakata and Kusunoki 1992). There is also a projection from MST to cells in the little studied posterior section of area STP (STPp) of the temporal lobe (see Figure 7.1). Cells in area STPp, like areas MT and MST, show selectivity for particular directions and types of motion. Most cells of area STPp

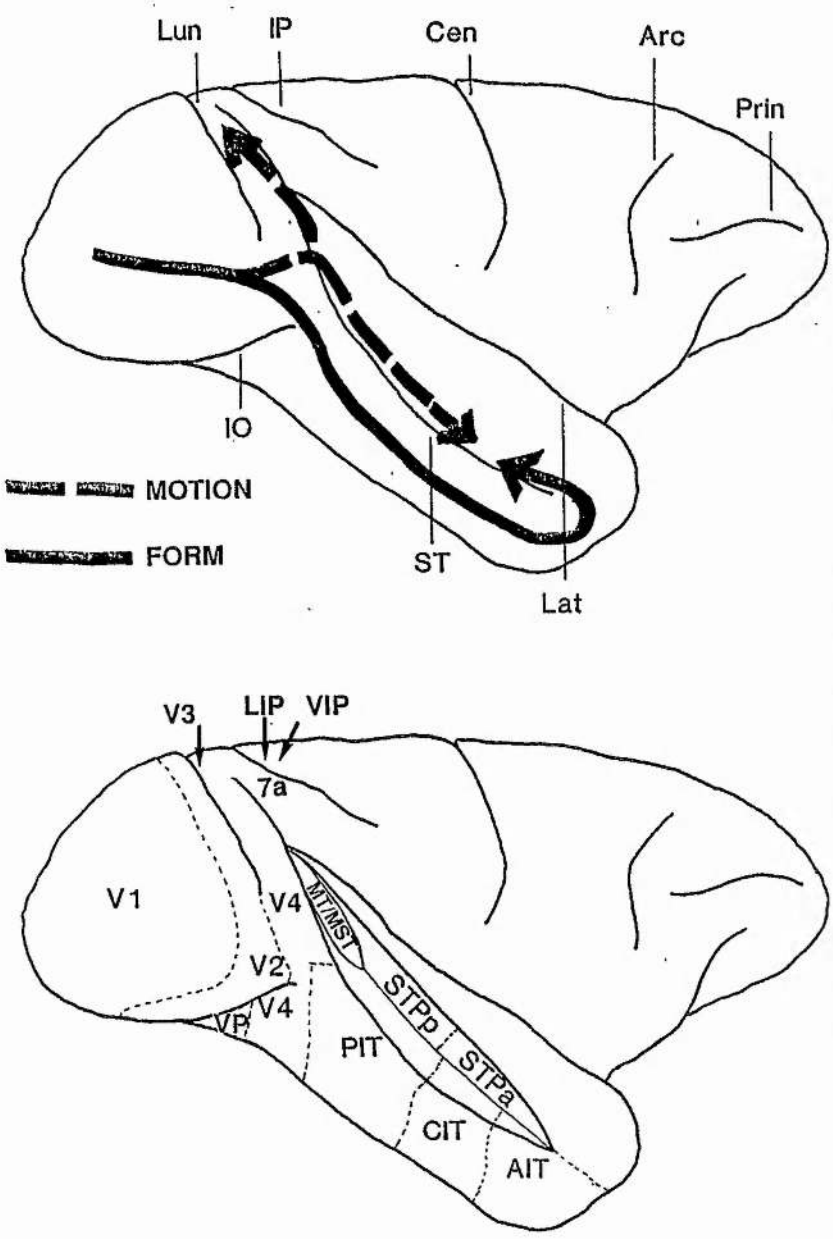


FIGURE 7.1. THE FORM AND MOTION PATHWAYS OF THE MACAQUE VISUAL SYSTEM. **UPPER:** Schematic diagram of the right view of the macaque brain showing the major sulci and the route of the form and motion pathways. [Abbreviations: Lun, lunate sulcus; IP, intra-parietal sulcus; Cen, central sulcus; Arc, arcuate sulcus; Prin, principle sulcus; IO, inferior occipital sulcus; ST, superior temporal sulcus; Lat, lateral sulcus]. **Lower:** The major cortical visual areas of the form and motion pathways. The superior temporal sulcus has been opened out to reveal the visual areas therein. Area V3 lies within the occipito-parietal sulcus, areas LIP and VIP lie within the intra-parietal sulcus. [Abbreviations: V1, V2, V3, V4, first second, third and fourth visual areas; VP, ventral occipto-parietal area; PIT, CIT, AIT posterior, central and anterior infero-temporal areas respectively; MT, middle temporal area; MST, medial superior temporal area; LIP, VIP, lateral and ventral intra-parietal areas; STPp, STPa, posterior and anterior superior polysensory areas.]

appear unlikely to signal the form of an object as their responses are determined by motion signals independent of the form of the moving object (Hikosaka et al. 1988), with the exception of "predictability" of the moving object (Hietanen & Perrett 1993).

The ventral pathway continues from V1/V2 to area V4 and then into the infero-temporal cortex (posterior, central and anterior areas, PIT, CIT and AIT of Figure 7.1). Area V4 has long been associated with the processing of colour information (Zeki 1973, 1977, 1980), although lesion experiments suggest its role in colour processing may not be as clear as once thought (Heywood and Cowey 1992; Heywood et al. 1988; Walsh et al. 1992a,b). Recent studies have also indicated that neural response selectivity within area V4 is related to the shape of the stimulus as well as colour (Van Essen and Zeki 1978; Gallant et al. 1992; Tanaka et al. 1991; Kobatake and Tanaka 1994). Interestingly the division between the magno- and parvo-cellular pathways thought to carry motion and form information respectively to the cortex does not seem as clear in V4 as once thought since both channels may convey information to V4 neurons (Ferrera et al. 1993, 1994). The selectivity of infero-temporal cortical cells shows an increase in complexity from posterior to anterior sections (Tanaka et al. 1991, 1993; Kobatake and Tanaka 1994; Tanaka 1993; for reviews see Perrett and Oram 1993; Oram and Perrett 1994). Neurons in IT respond to shape defined by motion, colour or contrast (Sary et al. 1993). The responsivity to shape defined by motion may not be surprising given the recent evidence suggesting that sensitivity to motion defined contours may be established as early as V1 in both humans and monkeys (Lamme et al. 1994).

In comparison with the dorsal stream of processing, selectivity for direction of motion is not so prevalent in the ventral stream (V4,IT), although there is some evidence that V4 may be involved in peripheral tasks requiring matching of direction (Maunsell 1994). In the anterior section of IT (AIT) cells selectively responsive to the face have been reported (Baylis et al. 1985, 1987;

Hasselmo et al. 1989a,b; Rolls 1984, 1992; Yamane et al. 1988; Young and Yamane 1992). Studies of face processing in humans using indwelling electrodes indicates complex pattern selectivity (e.g. faces, cars, words) is present in the fusiform and infero-temporal gyri (Allison et al. 1994). Thus, in both humans and macaque monkeys, the neurophysiological properties of the ventral pathway (V1-V2-V4-IT) suggest it is associated with complex form analysis, whereas the dorsal pathway is associated with the processing of motion and the co-ordination of motor output.

Neuropsychological evidence for two pathways

Lesion of areas MT and MST in the macaque produces deficits in motion perception while leaving form perception relatively uneffected (Newsome & Pare 1988; Cowey and Marcar 1992; Macar and Cowey 1992; Andersen and Siegel 1989; Siegel and Andersen 1986; Schiller 1993). In particular the evidence suggests that disambiguation of scenes with complex motion signals was severely impaired (Marcar and Cowey 1992). It should be noted that the segregation of scenes using motion cues is likely to be important for processing of the image motion components (Stoner and Albright 1993), but that this segregation does not in itself provide cues as to the identity of the moving object: only the location and orientation of the boundaries of (possibly multiple) objects within the scene. While this information would be important for establishing the object's identity, such form information is not explicit within any one MT/MST cell's response. Reaching and other visuo-spatial task deficits have been described following parietal lesions in monkey (e.g. Milner et al. 1977; Faugier-Grimaud et al. 1978; Gaffan and Harrison 1993). Deficits of motion perception have been documented for human patients with damage to the occipito-parietal cortex (e.g. Zihl et al. 1983; Vaina et al. 1990).

Lesion of macaque area V4 produces shape discrimination deficits (Heywood & Cowey 1987; Walsh et al. 1992a; Schiller 1993). Bilateral lesion of

monkey IT cortex also produces impairments in object recognition and in the learning of novel objects but not in many other visual tasks (Dean 1976; Weiskrantz and Saunders 1984). Lesion of IT also impairs recognition of shape defined by motion (Britten et al. 1992). The available data from human studies also indicates that lesion of the occipito-temporal pathway produces specific impairment of object recognition yet can leave acuity and colour discrimination relatively undamaged (Milner et al. 1991; Warrington 1986; Goodale and Milner 1992; Milner and Goodale 1993). The specific deficit of prosopagnosia (the inability to recognize familiar faces) in humans is also associated with damage to the occipito-temporal pathway (Sergent and Signoret 1992a,b).

The convergence of the two pathways

The major division in cortical visual processing of the macaque is therefore thought to be into the ventral pathway (V1-V2/3-V4-PIT-CIT-AIT) processing form or what an object is, and the dorsal pathway (V1-V2/3-MT-MST-STPp/parietal cortex) processing where an object is and its motion (Hubel & Livingston 1987; Ungerleider & Mishkin 1982; Felleman and Van Essen 1991). There is evidence to suggest that these pathways are separate from the retinal ganglion cells through the striate cortex, and indeed for most of the prestriate cortical visual areas (Felleman & Van Essen 1991; Young 1992), although this division might not be as clear as once thought (Ferrera et al. 1993, 1994; Merigan et al. 1991; Martin 1993). Evidence from studies of humans with local brain lesions suggests that this division of form and motion processing found in monkeys is also present in humans. Positron emission topography (PET) studies in normal humans have also indicated a separation of shape processing and motion processing cortical areas (Corbetta et al. 1990; Zeki et al. 1993; for review see Ungerleider & Haxby 1994).

Area STPa of the macaque receives input from area MST (in the dorsal pathway) via the posterior sections of area STP (STPp) and from AIT (in the

ventral pathway). It is thus one of the few visually responsive cortical areas where these pathways converge (Young 1992). One population of cells in area STPa that has been widely studied selectively responds to the form of an object (the head and body) independently of its motion (Gross et al. 1972; Bruce et al. 1981; Desimone et al. 1984; Perrett et al. 1982, 1985, 1987, 1988, 1991, 1992; Rolls et al. 1985, 1987, 1989; Baylis et al. 1985; Hasselmo et al. 1989a,b; Young & Yamane 1992). A second population of cells is selective for the direction of motion of objects, independently of their form (Bruce et al. 1981; Perrett et al. 1985; Hikosaka et al. 1988; Mistlin & Perrett 1990; Hietanen and Perrett 1993; Oram et al. 1993).

Here the response characteristics of a third population of cells in macaque area STPa which are selective for both the stimulus form (body view) and direction of motion (left, right, towards, away) are described. This cell population has been noted previously (Bruce et al. 1981; Perrett et al. 1985b, 1989a, 1990a,b), but has not been the subject of extensive study. In this chapter the classification of cell response types is given and evidence that the cell selectivity reflects binding of form and motion information is provided.

METHODS

Four rhesus macaque monkeys were used (*Macaca mulatta*, 3 male B, D, H wt. 5-8 kg, 1 female J wt. 7 kg from a UK registered breeding colony). The standard training, recording, behavioural task and histological reconstruction methods were used (see Chapter 2).

Stimuli

The stimuli were either real 3-D presentations or projected sequences of frames from a video disc. They included images of the experimenter walking, both forwards (compatible movement) and backwards (incompatible movement) in different directions (towards, away, left and right). Video images were taken of actors walking under strong diffuse lighting and subsequently stored on video disk. The luminance values of the video disk images were 0.2 cd/m^2 for the background and 4.0 cd/m^2 for the natural image of a walking person. Control objects moving in the same directions as the walking stimuli were used. These were matched for size and, like walking, exhibited articulated motion (e.g. patterned curtains), and were moved at the same speed (+/- 15%) and direction (+/- 10 degrees) as the walking stimuli. Control object motion was also recorded and stored on video disk. Video disk images were projected to be life-sized.

Stimuli were also created using a computer-video system (Fairlight) that allowed a textured background to move with a superimposed image of the head remained stationary. Conversely, the video image of the head could be moved while the background remained stationary. The background consisted of black discs (subtending 1-10 degrees) on a white surface. Front and back views of the head static, looming and receding were combined with static, looming and receding background. Likewise, left and right profile views of the head that were either static or moving left and right were combined with a static background and leftward and rightward background motion. These images were stored on video-disk, and could be replayed in pseudo-random order under computer control.

Stimulus presentation

The minimum testing protocol consisted of two opposite body views and controls moving in two opposite directions (e.g. left profile moving left and right, right profile moving left and right and control objects moving left and right). The view and direction combinations were divided into two categories: (1) Incompatible movement where the image was of someone walking backwards and

(2) Compatible movement which was defined as walking forwards (following one's nose).

Cells found to be responsive to moving bodies were also tested for selectivity to the single limb movements present in the preferred stimulus. These tests involved the presentation of the arm or leg flexing and extending in isolation (i.e. rest of the body occluded from sight or visible but stationary). Cells found selective for these stimuli have been reported previously (e.g. leg, arm or hand motion, see Perrett et al. 1985b, 1989a,b, 1990a,b, 1993; Mistlin and Perrett 1990). The cells reported here, however, were unresponsive to individual limb movements.

Data analysis

A cell was provisionally classified as selective for the sight of body movement if there was a significant overall effect of conditions and one direction / body view combination was different ($p < 0.05$) from (i) a second body view moving in the same direction and (ii) the same body view moving in a second direction. Off line analysis for all cells took the form of 2-way ANOVA, with the direction of motion as one factor and the stimulus type (compatible, incompatible and control object) as the second factor. A second 2-way ANOVA was performed with body view as one factor and stimulus type (compatible, incompatible, static) as the second. Significance for all statistical tests was taken at the 0.05 level. Post hoc testing of the ANOVAs was performed using the protected least significant difference (PLSD) test (Snedecor & Cochran, 1980). A cell was classified as being selective for body motion stimuli using the results of the 2-way ANOVA analyses where the post-hoc testing confirmed the preliminary classification (see above).

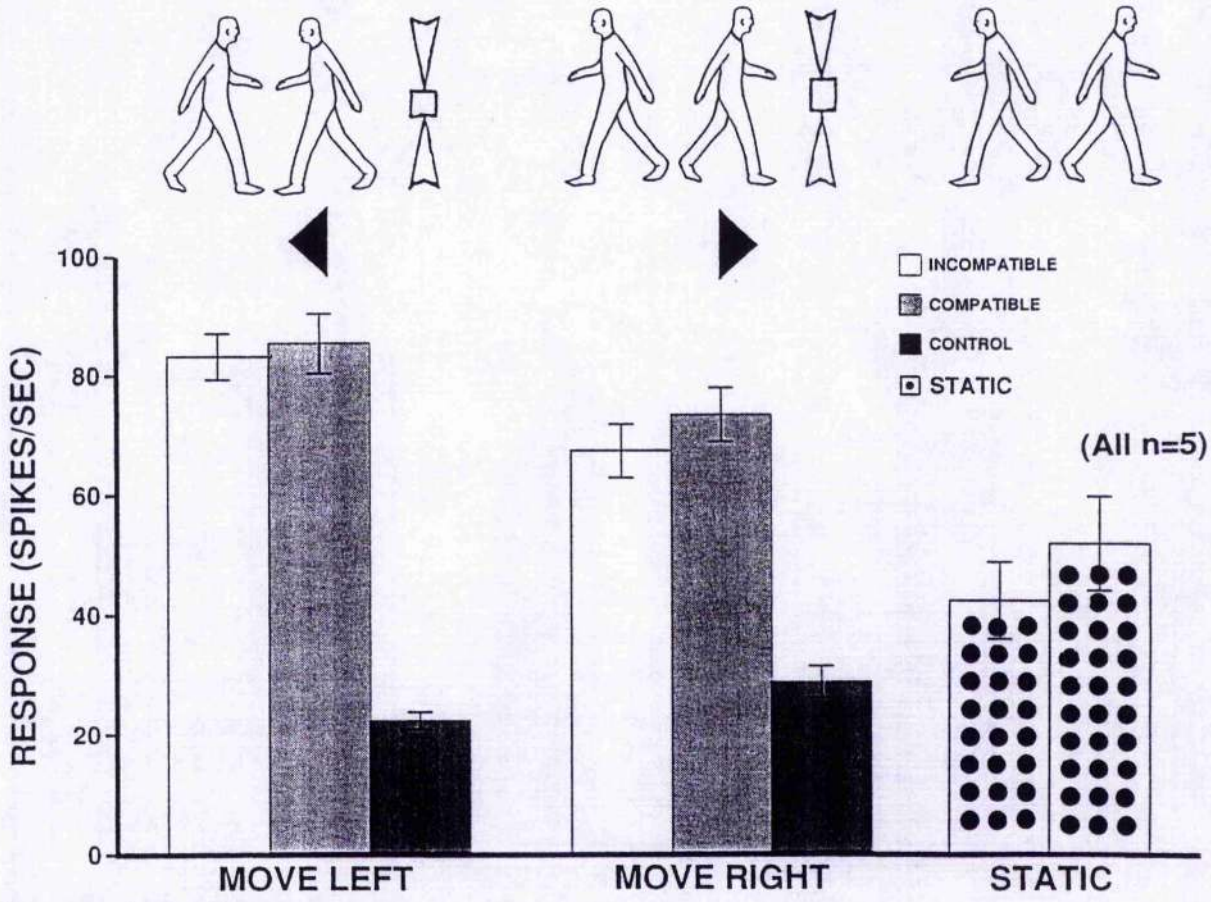


FIGURE 7.2. CELL SELECTIVITY FOR MOVING BODIES. Mean response magnitudes (+/- 1 S.E.M.) for one cell to bodies facing left and right presented walking to the left and to the right or as a static image. Schematic representations of the stimuli are presented above the histogram bars. The cell responded more strongly to bodies moving left or right regardless of the body view than the response to controls moving in the same directions ($p < 0.0005$ each comparison). The response to static presentation of the same bodies views produced a response that was weaker than the response to the moving images of the body ($p < 0.005$ each comparison). [Overall effect of moving stimuli: Effect of direction, $F_{[1,44]} = 7.1$, $p = 0.01$, Effect of motion type (compatible, incompatible, control) $F_{[2,44]} = 63.6$, $p < 0.0005$, interaction $F_{[2,44]} = 2.5$, $p > 0.05$. Overall effect of ANOVA by view: Effect of view, $F_{[1,44]} = 0.7$, $p > 0.4$; Effect of motion (compatible, incompatible, static) $F_{[2,44]} = 17.5$, $p < 0.0005$; interaction $F_{[2,44]} = 4.8$, $p < 0.05$.]

RESULTS

From the four experimental subjects a total of 6,459 cells were screened for visual responses. Of these, 1507 (23%) were found to be responsive to visual stimuli. 348 cells were classified as showing sensitivity to the form and also to the motion of the stimulus. These cells included sensitivity to single limb motion, body rotation and hand manipulations. A total of 161 cells (11% of visually responsive cells) showed that both body view (form) information and direction of motion (towards, away, left or right) information was present in their response characteristics. This latter cell population is the subject of this chapter.

Partial selectivity for form and motion

We describe first cell response characteristics that indicate conjoint coding of form and motion, but where the selectivity is relatively low. Cells "partially selective" for form and motion gave stronger responses to moving bodies than to moving controls indicating some form specificity. Further the responses were also enhanced compared with the responses to the static images of the body indicating motion specificity. These cells were divided into those which responded to both compatible and incompatible motion of the body (moving body selective cells) and cells responsive to the motion of only one body view (moving body-view cells).

(a) Moving body selective cells

A total of 10 cells showed significantly larger responses to moving bodies than to control objects moving in the same direction and static images of the body. Figure 7.2 shows an example of such a cell, where the mean response from 10 trials is plotted (+/- 1 S.E.M.). It can be seen that for this cell the response to bodies moving to the left or right was greater than the response elicited by control objects (matched for size, non-rigid motion, see methods) moving in the same

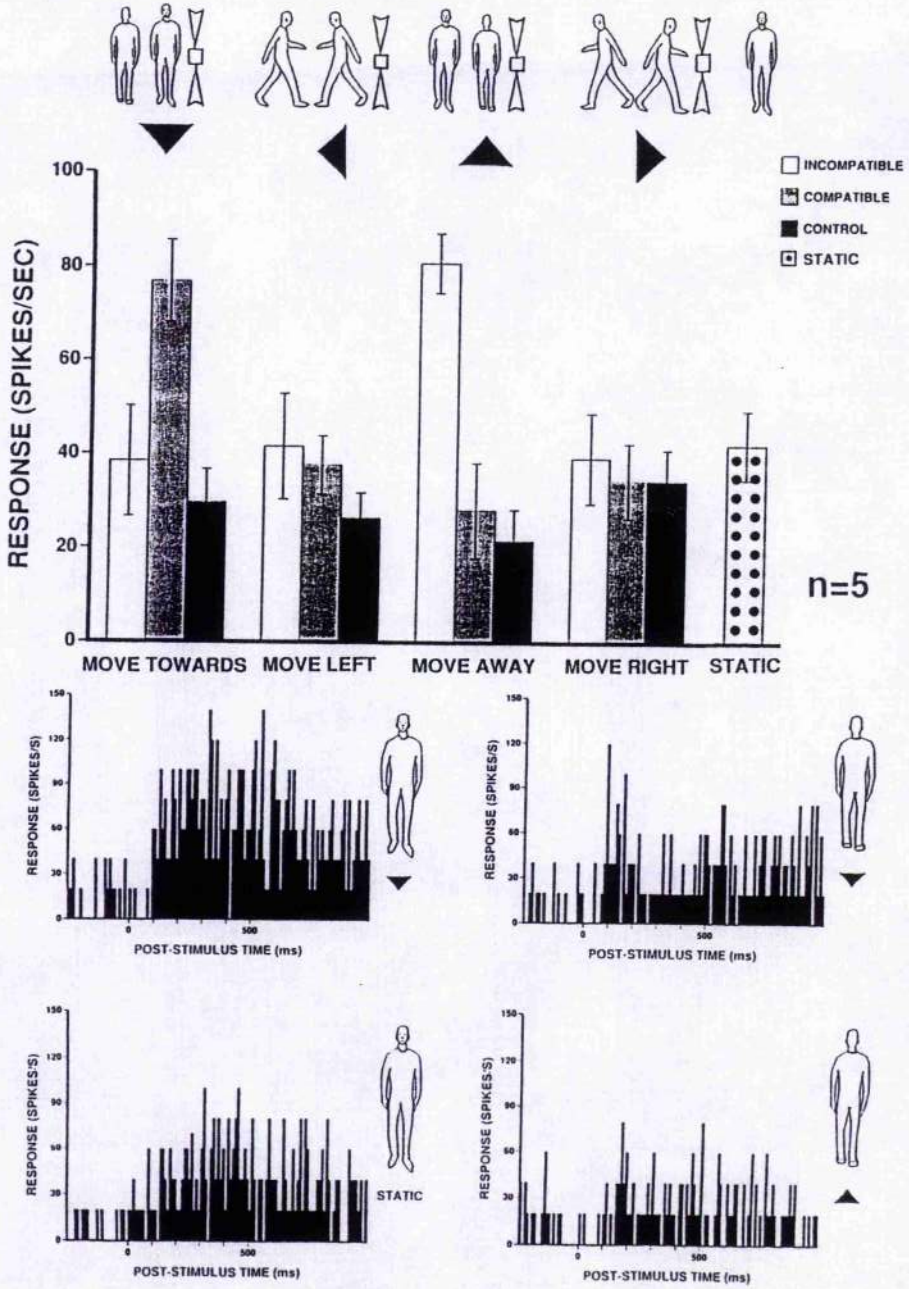


FIGURE 7.3. CELL SELECTIVITY FOR MOVEMENT OF ONE BODY VIEW. UPPER: Histogram of mean responses (+/- 1 S.E.M.) to compatible, incompatible and controls moving towards, left, right and away from the monkey, and static face view. The stimuli are represented schematically above each histogram bar. The cell responded more strongly to moving face view bodies than to control objects or the back view of the body moving in the same directions or to other body views moving in other directions ($p < 0.0005$ each comparison). the response was also stronger when the face view of the body was moving than when it was stationary ($p < 0.0005$ each comparison). [Overall effect of moving stimuli: Effect of direction, $F_{[1,48]} = 1.7$, $p > 0.1$, Effect of motion type (compatible, incompatible, control) $F_{[2,48]} = 7.3$, $p = 0.002$, interaction $F_{[6,48]} = 5.8$, $p < 0.0005$.] **LOWER:** Peri-stimulus time histograms (PSTH) for compatible and incompatible movement towards and away from the subject of the cell illustrated in the upper section. The stimuli are represented to the right of each PSTH.

directions. The cell, however, did not respond differentially to the two different body views tested (compatible and incompatible). There was for this cell some slight directional selectivity, with movement to the left producing a small but significantly larger response than movement to the right. The response to moving bodies was larger than the response to the same body view when static.

(b) Moving body-view selective cells

A further 13 cells were found that showed selectivity for a particular body view when moving, but showed no selectivity for the direction of motion. Figure 7.3 shows an example of a cell with this type of response selectivity. For this cell, a large response was seen when the front body view moved either towards or away from the monkey. The back view of the body (and other control objects) moving in these same directions produced a significantly smaller responses. Static images of the same body view did not elicit as large a response. As can be seen in the lower part of Figure 7.3, the cell showed small responses to the other tested stimuli, including presentation of the static preferred body view. These responses were small, however, and did not show view discrimination (not shown).

Assessment of eye position and eye movements

The behavioural testing paradigm used allowed the subjects to move their eyes during stimulus presentation. As noted in Chapter 2, monitoring of eye position allowed assessment of the velocity of eye movement as well as the eye position. No systematic relationship between either eye position or eye velocity and cell response was observed for any of the cells reported here. Figure 7.4 shows rastergram displays of the responses from ten trials of one cell to four stimulus types. The left block of the Figure (7.4a-d) shows the horizontal and vertical eye position traces, the right block (7.4e-h) shows the horizontal and vertical eye velocity traces for each of the trials. The cell responded strongly to the sight of incompatible walking when a body moved away from the subject

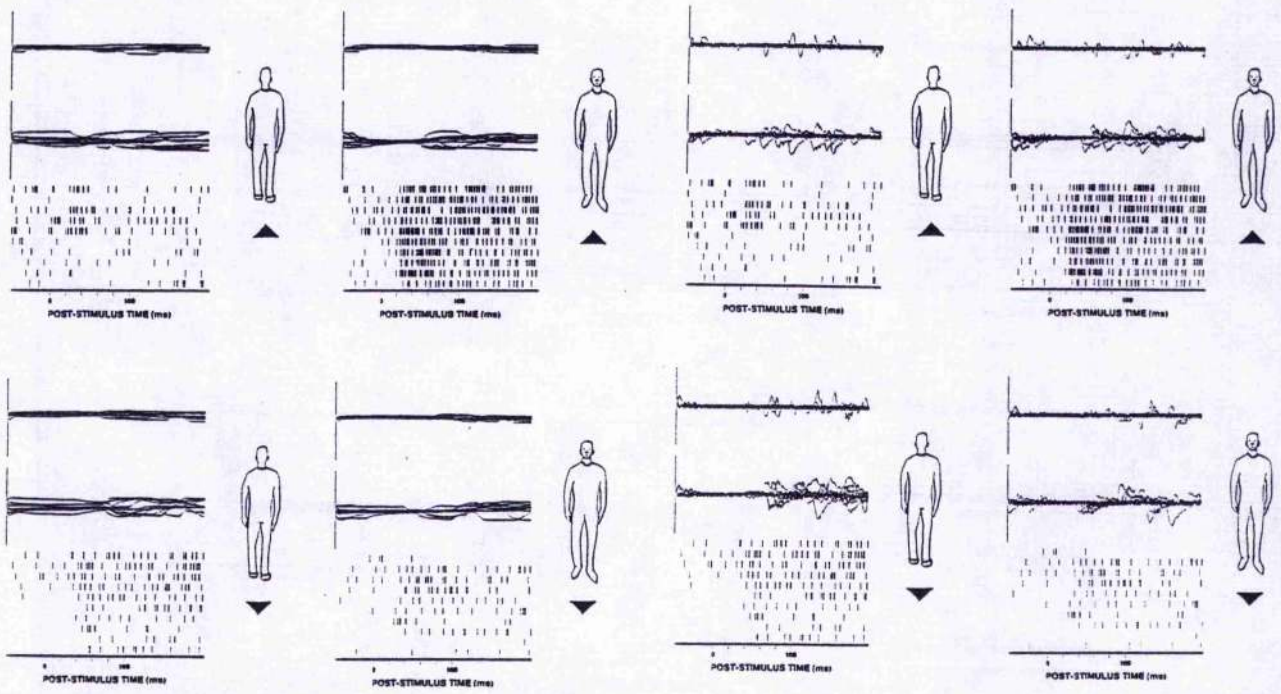


FIGURE 7.4. EYE POSITION AND MOVEMENT VELOCITY CANNOT ACCOUNT FOR UNIMODAL SELECTIVITY FOR INCOMPATIBLE WALKING. A-D) Horizontal and vertical eye position traces for a cell showing selectivity for incompatible walking away from the monkey. Full scale deflection = 50 degrees. The rastergrams show the cell response for each of ten trials under the stimulus condition represented schematically to the right of each Figure. A) compatible walking away from the monkey, B) incompatible movement away. C) incompatible movement towards the monkey. D) compatible movement towards. E-H) Horizontal and vertical eye velocity traces and rastergrams for the same cell as shown in a-d. Full scale deflection = 500 degrees/sec. E) compatible walking away from the monkey, F) incompatible movement away. G) incompatible movement towards the monkey, H) compatible movement towards.

(front view move away, 7.4b,f). In all four stimulus conditions (including the most effective) the horizontal eye position showed little variation up to some 250-300 ms post-stimulus onset (7.4a-d, upper traces). For the effective stimulus (7.4b) vertical eye position traces indicate that fixation was also maintained for nearly 250 ms post-stimulus onset.

The time at which fixation of the coloured LED was broken is best illustrated in Figure 7.4, parts e through to h. Here onset of eye movement is marked by an upward or downward deflection in the eye velocity traces. Figure 7.4f shows the horizontal and vertical eye velocity traces during presentation of the preferred stimulus. As is evident, the responses are tightly time locked to the visual stimulation and occur 100-150 ms post-stimulus onset, whereas the saccades occur from 150-1000 ms post-stimulus onset. Furthermore, comparison of Figures 7.4e and f indicates that saccade velocity, frequency and timing are comparable between these two conditions, yet there is a clear difference between responses. In Figures 4.4 g and h the number of saccades is higher than in Figure 7.4 e and f, yet again the responses are clearly reduced compared to the preferred stimulus (7.4f). Therefore it is highly unlikely that the neural responses are related to onset of saccades.

In summary, examination of eye position and eye velocity was used to assess the possibility that the observed cell responses were related to eye movements. While one cannot rule out a relationship completely (see also discussion), the evidence suggests that, for each of the cells reported here, there is considerable overlap in the type of eye movements occurring during effective and ineffective stimulus conditions. Despite this overlap there is a clear response difference related to the stimulus condition. In remaining figures either the horizontal eye position (when stimulus movement was to the left or right) or the vertical eye position traces (when stimulus movement was towards or away) are displayed. This choice of display provides an indication of the similarity of eye movements between stimulus conditions.

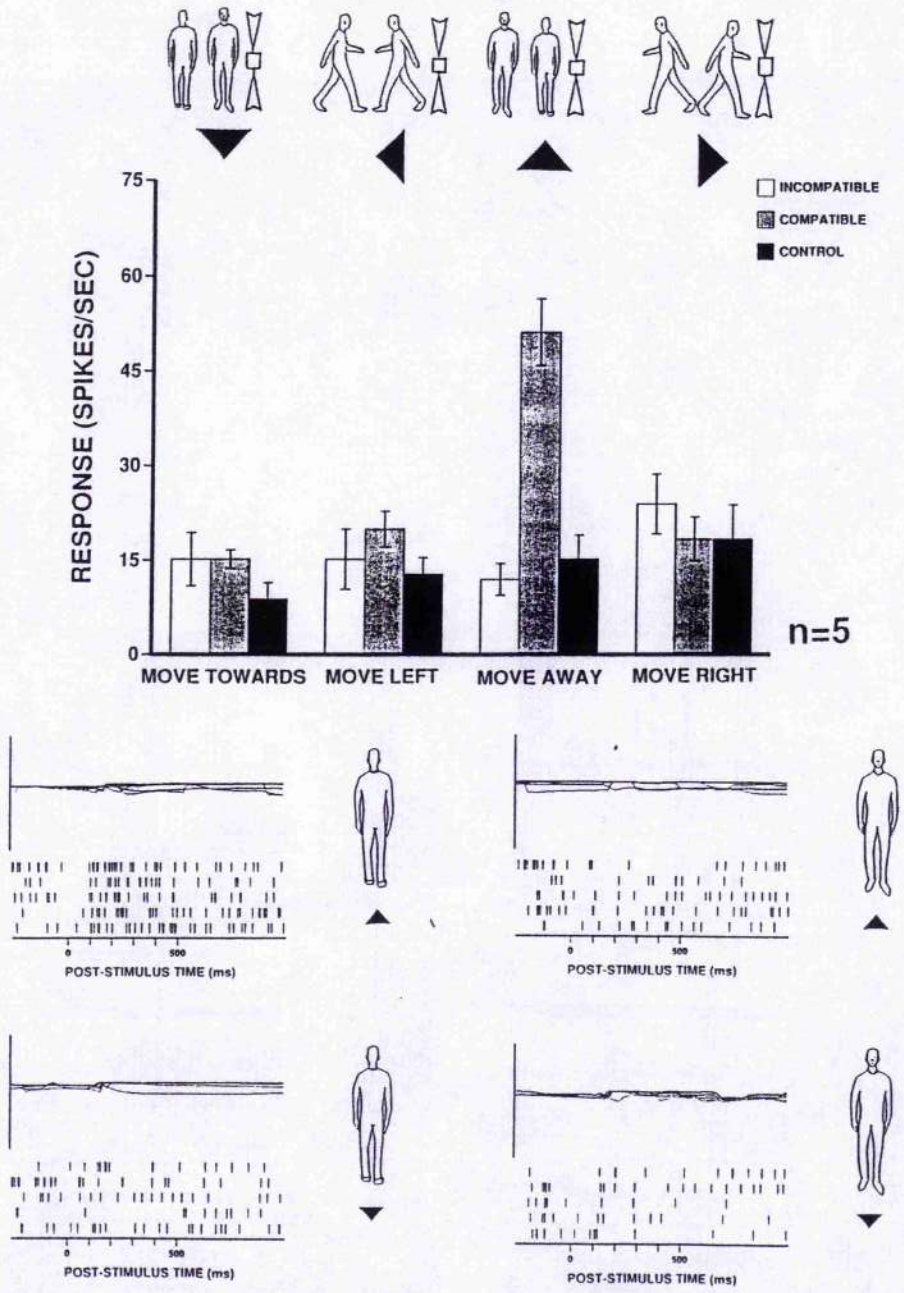


FIGURE 7.5. RESPONSES INDICATING THE INTEGRATION OF FORM AND MOTION AT THE SINGLE CELL LEVEL. **UPPER:** Response magnitudes (+/- 1 S.E.M.) from a single cell to walking bodies and controls moving in 4 directions. The cell gave a greater response to the back view of the body walking away (compatible walking) from the monkey than to either controls or the front view of the body moving away from the monkey ($p < 0.0005$ each comparison). The cell did not respond to other body views moving in other directions ($p < 0.0005$ each comparison), or to static images of the body (not shown). [Overall effect of conditions: Direction of motion $F_{[3,48]} = 6.5$, $p = 0.001$, Mode of motion (compatible, incompatible, control) $F_{[2,48]} = 11.3$, $p < 0.0005$, Interaction $F_{[6,48]} = 7.6$, $p < 0.0005$]. **LOWER:** Vertical eye position traces and individual rastergram displays of the cell response to compatible and incompatible movement towards and away from the monkey. The similarity of the eye positions and of the magnitude of small saccades between stimulus conditions rules out the possibility of the cell response being related to eye movements.

Cell coding of specific view and direction combinations

A total of 138 of the 161 studied cells were found to be sensitive to a particular combination of body view (form) and direction of motion (see methods). The majority of cells sensitive to both form and motion displayed a pattern of responses that indicated a high selectivity for both body view and direction of motion. These cells could be divided into four categories.

(a) Unimodal coding of form and motion

Cell that responded strongly to only one of the view and direction combinations tested were classified as unimodal. Figure 7.5 shows an example of such a cell. The upper section shows the mean response magnitudes to incompatible (walking backwards), compatible (walking forwards), and control objects moving in four directions. Note that the response cannot be explained as sensitivity to body view, since the preferred body view (back view) moving towards the monkey produces a weaker response than the preferred view moving in the preferred direction (away). Likewise, the response selectivity cannot be explained in terms of simple direction preference, since the front view and control objects moving away from the monkey did not produce a response. We refer to cells with this type of response (responding to only one view and direction combination) as "unimodal". The majority of the cells sensitive to view and direction combinations were unimodal (125/138, 90.5%).

Vertical eye position traces and individual trial rastergrams are shown in the lower section of Figure 7.5. The top left section shows the individual responses and the eye position to the preferred view and direction combination. It can be seen that the monkey saccades upwards on most trials. These eye movements occur approximately 100 ms after response onset. Under other stimulus conditions the same saccades are made, yet there is no clear cell response. Further, during presentation of the preferred stimulus condition the

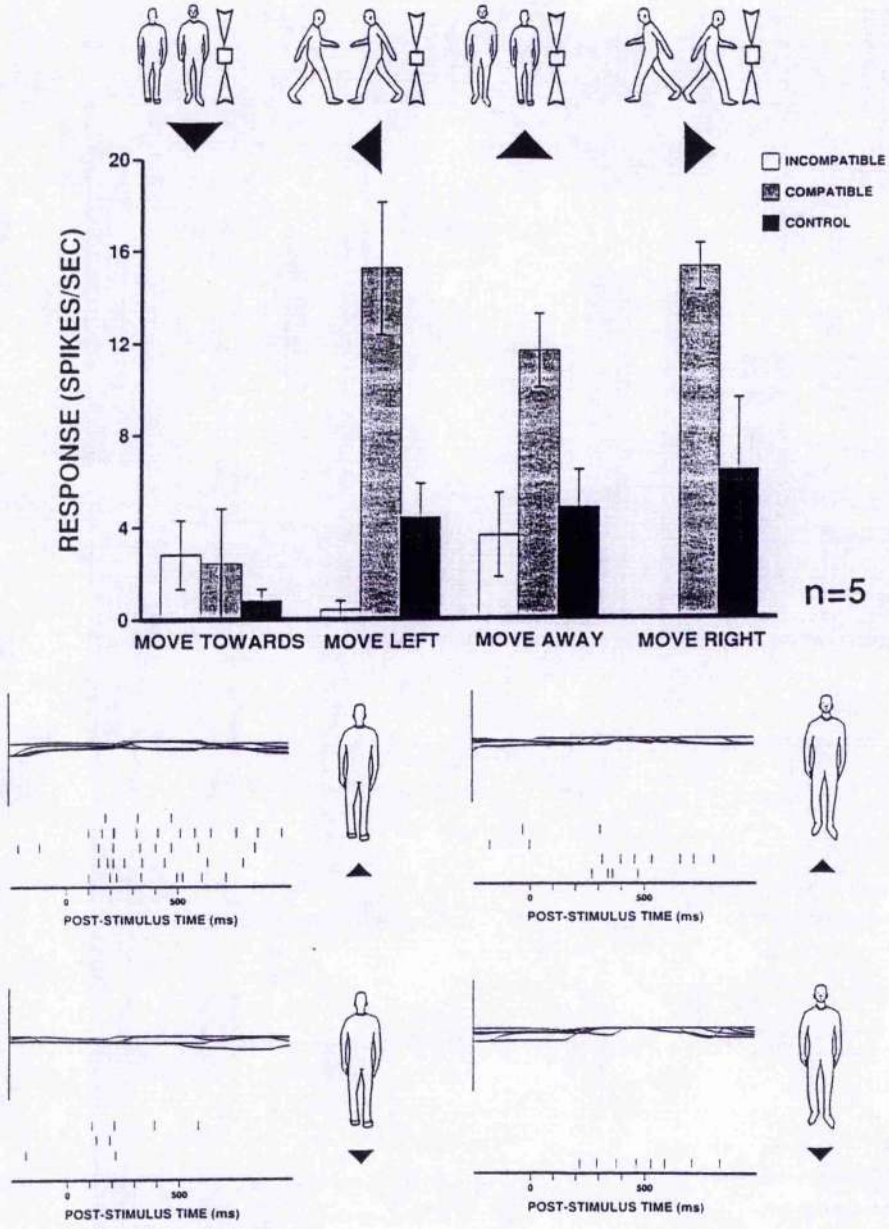


FIGURE 7.6. BROAD TUNING OF FORM AND DIRECTION SELECTIVITY. **UPPER:** Response magnitudes (+/- 1 S.E.M.) of a single cell selective for compatible walking over a broad range of directions. The cell response to compatible walking to the left, right or away was greater than the responses of control objects or incompatible walking in the same directions ($p < 0.0005$ each comparison). [Overall effect of conditions: Direction of motion $F_{[3,48]} = 5.6$, $p < 0.005$, Mode of motion (Compatible, incompatible, control) $F_{[2,48]} = 30.0$, $p < 0.0005$, Interaction $F_{[6,48]} = 4.3$, $p = 0.001$]. **LOWER:** Vertical eye position traces and individual rastergram displays of the cell response to compatible and incompatible movement towards and away from the monkey. The considerable overlap between eye position traces and the similarity of the saccades between stimulus conditions suggests that the responses are not due to differential eye movements.

response latency is 100 ms post-stimulus onset in all trials. In contrast the time of the upward saccades varies from 180-550 ms post-stimulus onset. Again, the differential cell responses under the various stimulus conditions therefore cannot be correlated with the eye movements.

(b) Broad coding of form and motion

A small number of cells (7/125, 6%) that were included in the unimodal category showed relatively broad tuning. Figure 7.6 shows an example of a cell showing broad tuning of form and direction. Note that while selectivity for compatible compared to incompatible movement was high for movements left right and away, the cell responded to compatible walking in all three of these directions.

The lower section of Figure 7.6 shows the vertical eye position trace and rastergrams during presentation of compatible and incompatible movement towards and away from the subject. There are no systematic differences between the eye positions during the preferred combination (back body view retreat) and the other stimulus conditions. For instance, when the back body view walked either away from or towards the subject, on two trials the subject maintained fixation of the coloured LED, whereas for the remaining three trials the subject saccaded (upward to the head of the walking body), yet despite the similarity of eye position and movements the difference in response magnitude across conditions is clearly visible from the rastergram displays.

Cells were classified as broadly coding form and motion if two adjacent directions of movement (e.g. walking left and walking away) for a particular type (e.g. compatible) of walking showed higher responses than controls or the opposite body view (e.g. incompatible walking) moving in the same directions.

(c) Bimodal coding of form and motion

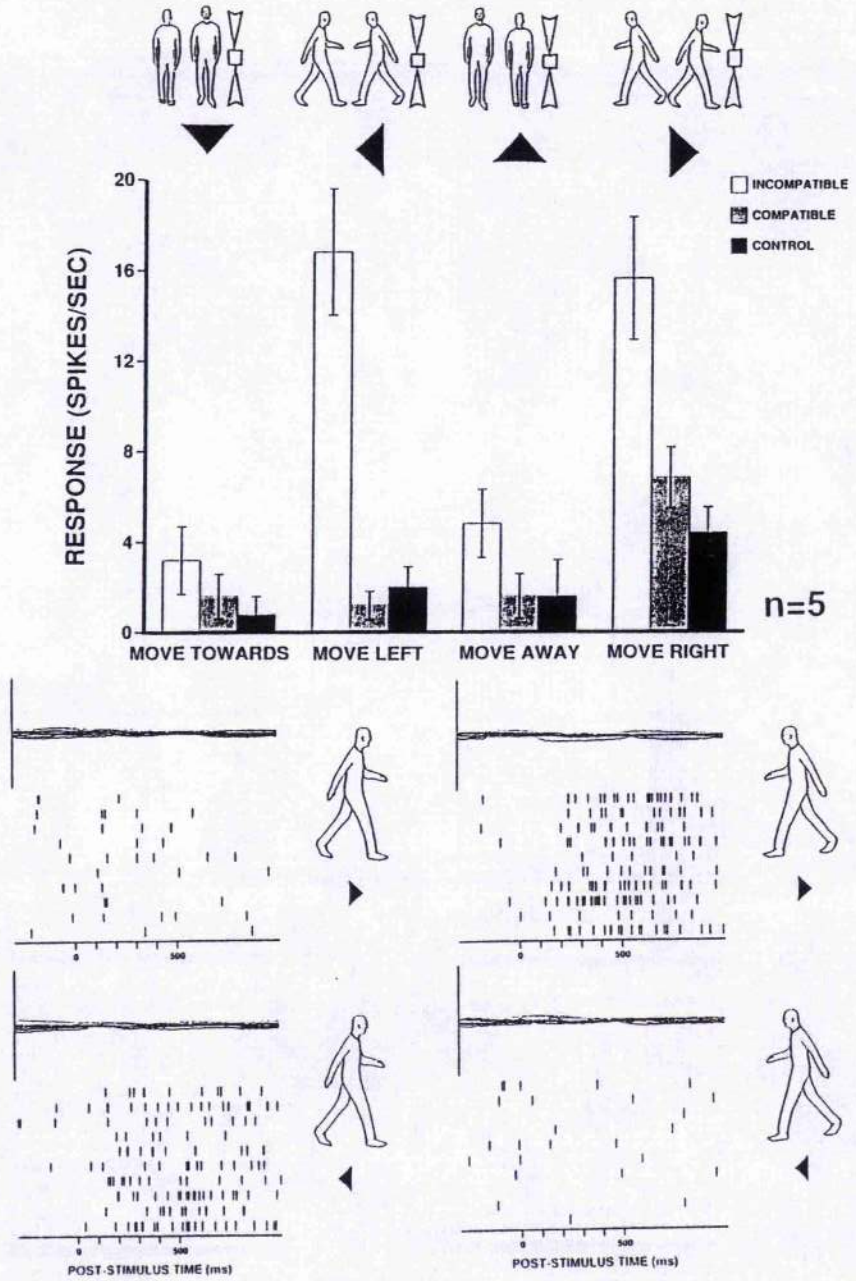


FIGURE 7.7. CELL RESPONSES SHOWING A BIMODAL SELECTIVITY FOR FORM AND DIRECTION COMBINATION. **UPPER:** Response magnitudes (+/- 1 S.E.M.) of one cell to walking bodies and controls showed a preference for incompatible walking to the left or to the right. Both the left profile view walking to the right and the right profile view walking to the left produced responses above compatible walking or control movement in the same directions ($p < 0.0005$ each comparison). The cell did not respond to bodies or controls moving either towards or away from the monkey ($p < 0.0005$ each comparison). [Overall effect of conditions: Direction of motion $F_{[3,78]} = 9.4$, $p < 0.0005$, Mode of motion (Compatible, incompatible, control) $F_{[2,78]} = 35.0$, $p < 0.0005$, Interaction $F_{[6,78]} = 3.8$, $p = 0.002$]. **LOWER:** Horizontal eye position traces and individual rastergram displays of the cell response to compatible and incompatible movement left and right. The similarity of the eye positions and of the magnitude of small saccades between stimulus conditions suggests that eye movements have no discernable effect on the cell response.

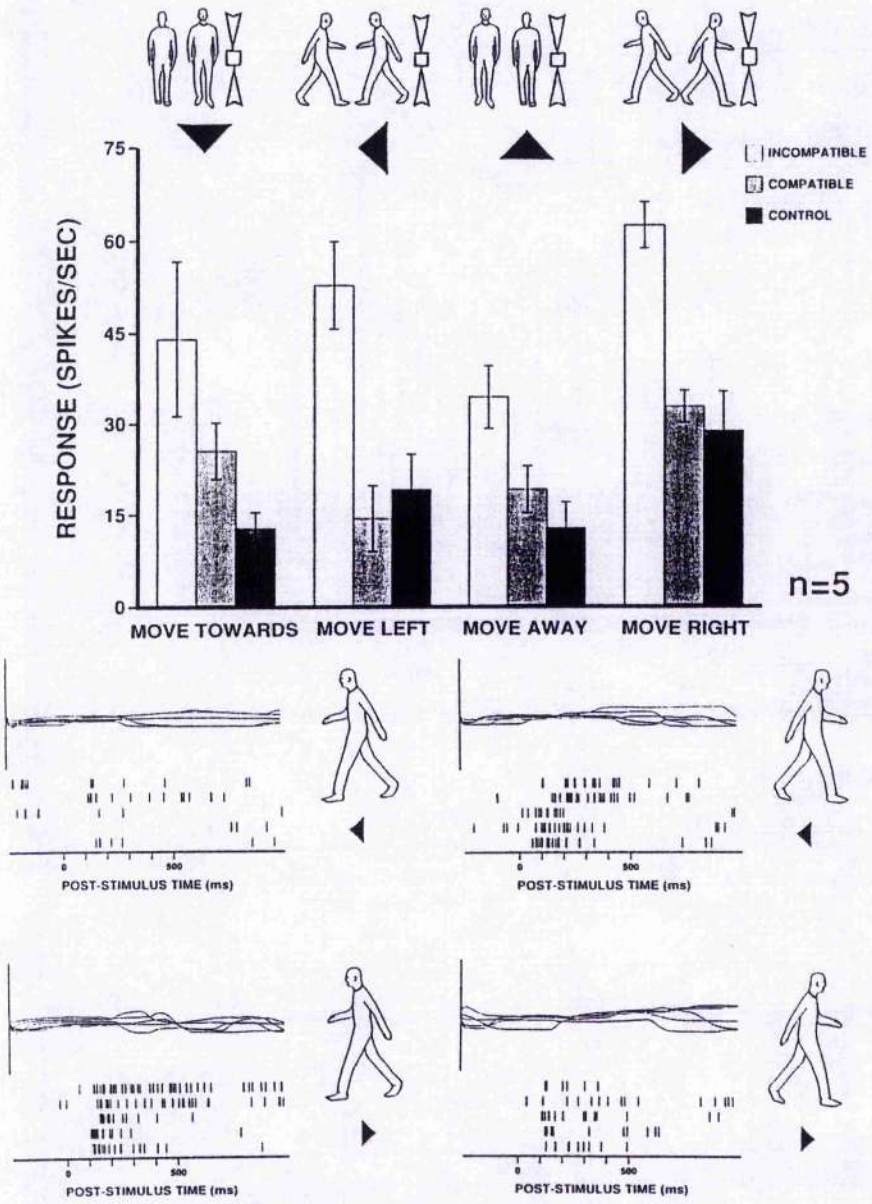


FIGURE 7.8. CELL RESPONSES INDICATING OBJECT CENTRED SELECTIVITY FOR INCOMPATIBLE WALKING. **UPPER:** Response magnitudes (+/- 1 S.E.M.) of a single cell to incompatible walking bodies was greater than bodies walking compatibly or controls moving in the same directions ($p < 0.01$ each comparison). [Overall effect of conditions: Direction of motion $F_{[3,48]} = 5.6$, $p = 0.002$, Mode of motion $F_{[2,48]} = 29.3$, $p < 0.0005$, Interaction $F_{[6,48]} = 0.9$, $p > 0.4$)]. **LOWER:** Horizontal eye position traces and individual rastergram displays of the cell response to compatible and incompatible movement left and right. The similarity of the eye positions and of the magnitude of the saccades between stimulus conditions indicates that the cell response is not related to eye movements.

A further 9 cells (6.5%) were found that responded to two non-adjacent combinations of particular view and direction combinations. Figure 7.7 shows an example of a cell showing a bimodal response profile. For this cell incompatible movement, either to the left or the right produced a response that was greater than the responses to controls or other body views moving in the same directions. Further, the cell did not respond to control or body movements either towards or away from the subject. Thus, cells were classified as showing bimodal coding if they responded to one type of walking (e.g. incompatible) in two directions of motion that were separated by directions for which the cell was unresponsive.

Horizontal eye position and rastergrams for each of the ten trials during compatible and incompatible walking to the left and to the right are shown in the lower section of Figure 7.7. As can be seen, there is considerable overlap in the position and movements (not shown) of the eyes for the different stimulus conditions, yet the responses occur only when the stimulus was incompatible walking to the left or right.

(d) Object centred coding of form and motion

A fourth cell response pattern found was for either compatible or incompatible walking independent of the direction of motion. We refer to this cell response pattern object-centered, since the response is related to the body's own axis (e.g. following the nose) and hence independent of the observer's viewpoint. Only 4 cells (3%) were found to show object centred response patterns. Figure 7.8 shows an example of a cell showing an object centred response pattern where incompatible walking in any direction produced a response greater than compatible walking or controls moving in the same directions.

Although response magnitude varied with direction (main effect of direction was significant, $p = 0.002$), the non-significant interaction term indicated that there was maintained selectivity for incompatible walking across all directions. The lower section of Figure 7.8 shows the horizontal eye position trace

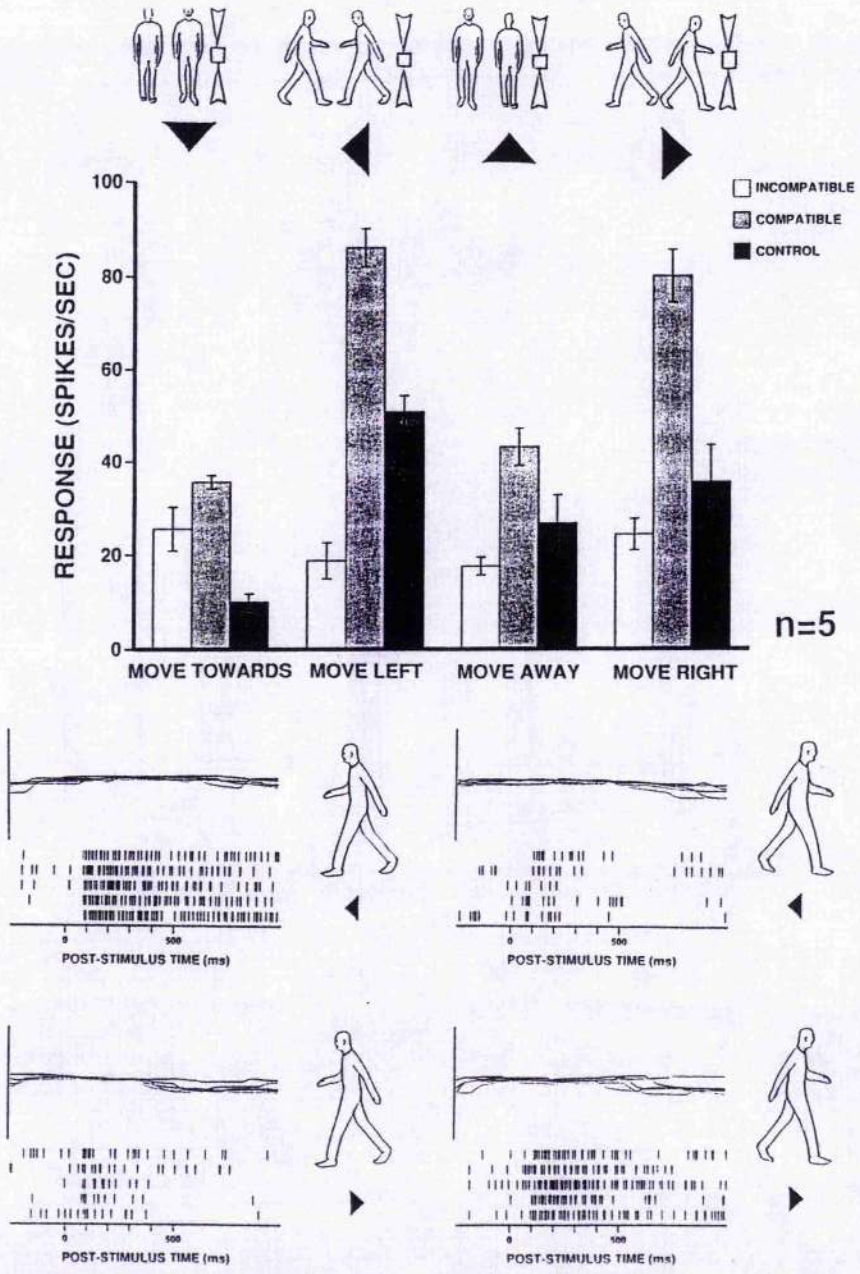


FIGURE 7.9. CELL RESPONSES SHOWING BOTH BIMODAL AND OBJECT-CENTERED SELECTIVITY FOR COMPATIBLE WALKING. **UPPER:** Response magnitudes (+/- 1 S.E.M.) to compatible, incompatible and control object movement in four directions. Although compatible movement in any of the directions produces a higher response than incompatible or control movement in the same direction ($p < 0.05$ each comparison), there is clearly a bimodal pattern of activity present, with compatible movement to the left and right producing larger responses than movement towards or away ($p < 0.0005$ each comparison). [Overall effect of conditions: Direction of motion $F_{[3,48]} = 28.2$, $p < 0.0005$, Mode of motion (Compatible, incompatible, control) $F_{[2,48]} = 88.3$, $p < 0.0005$, Interaction $F_{[6,48]} = 10.5$, $p < 0.0005$]. **LOWER:** Horizontal eye position traces and individual rastergram displays of the cell response to compatible and incompatible movement left and right. The similarity of the eye positions and of the magnitude of the saccades between stimulus conditions suggests that eye movements have little if any effect on the cell response.

and rastergrams of walking movements to the left and right. The overlap between stimulus conditions in both eye position and saccade size makes it unlikely that the response selectivity reflects differential eye movement patterns.

(e) Mixed form and motion coding selectivity

A small number of cells that showed object-centered response properties also showed sensitivity to the particular view/direction combination. Figure 7.9 gives an example of such a cell. While analysis indicated that this cell showed object-centered characteristics with compatible walking producing a stronger response than incompatible walking and controls in all directions ($p < 0.05$ each comparison), there was clearly a tendency for a strong bimodal response pattern.

The lower part of Figure 7.9 shows the horizontal eye position and rastergrams from the trials with compatible and incompatible movement to the left and right. Fixation was maintained for 200-500 ms post-stimulus onset under all stimulus conditions, yet responses latencies were tightly locked to 100 ms and large responses only occurred for compatible movements.

Coding of compatible and incompatible walking

As stated above, the majority of the cells sensitive to view and direction combinations were unimodal (125/138, 90.5%). This classification was assessed using the selectivity of the neurons to different combinations of body views and directions of motion. This classification system, however, gives little indication of the selectivity for the type of walking (compatible or incompatible). Table 7.1 gives a break down of the number of cells found in each of the above response categories. As can be seen, there were approximately twice as many cells sensitive to compatible walking as there were to incompatible walking.

Table 7.1

Class	Comp	Incomp	Total
Uni.	98	27	125 (90.5%)
(Br.	6	1	7)
Bi.	5	4	9 (6.5%)
O.C.	3	1	4 (3.0%)
(Mix.	2	0	2)
TOTAL	**106**	**32**	**138**

Table 1. Number of cells showing compatible or incompatible response selectivities to form and motion stimuli. (Comp = compatible, Incomp = incompatible; Uni = Unimodal, Bi. = bimodal, Br = unimodal with broad tuning, O.C. = object-centered responses independent of observer's vantage point, Mix. = object centred with bimodal properties).

Cell responses in area STPa show binding of form and motion information

The data presented so far indicates that cells in area STPa show clear response selectivity for particular combinations of body view (form) and direction of motion. The coding of information about an object's presence and that the stimulus also contains motion information is not, however, sufficient to explain the every-day experience of knowing *what* object is moving. For this to be signalled by single cells, the response selectivity should be such the juxtaposition of appropriate form signal from one object and motion signal from a second object does not activate the cell. In other words, the form and motion signals to which the cell responds should be bound together so that only one object is the

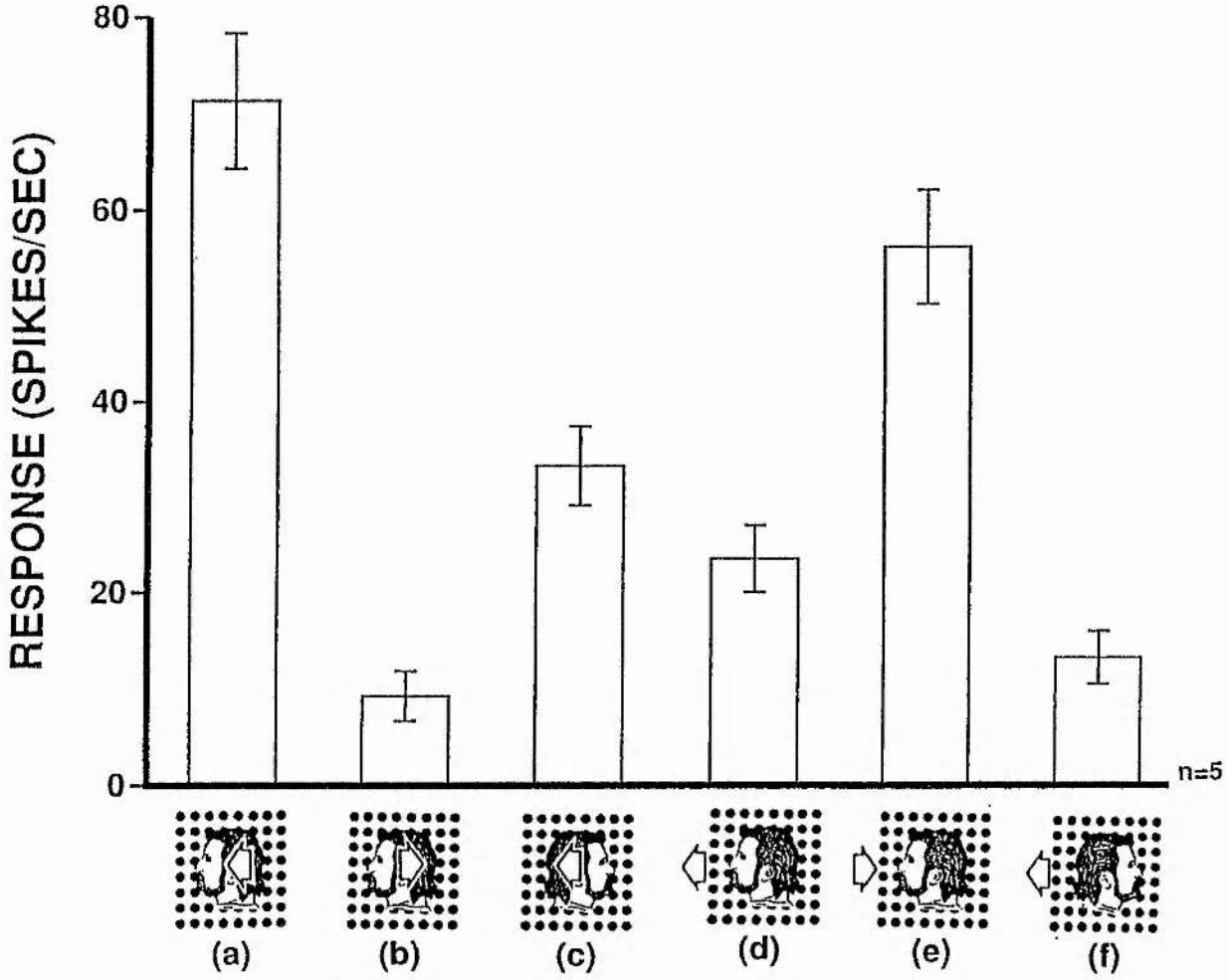


FIGURE 7.10. CELL RESPONSES INDICATING 'BINDING' OF FORM AND MOTION. Response magnitudes (+/- 1 S.E.M.) of cell to compatible head (and body) movement to the left. With a static background (a-c) the response to left profile moving left (a) was greater than other view direction combinations (b, c). With static head views and a moving background (d-f) the *same relative motion* (i.e. a rightward moving background and left profile view (e)) evoked larger responses than other view/direction combinations. Comparison of conditions (a) and (d) indicates the cell requires the correct relative motion of the face contours and not simply the presence of these contours combined with the preferred motion direction arising from other stimuli. This shows binding of form and motion information without sensitivity to the illusory conjunction present in (d).

source of the appropriate form and motion signals. This can be referred to as the "binding" of form and motion.

Figure 7.10 shows an example of the responses of a cell when tested with stimuli designed to investigate the "binding" of form and motion information. During testing in which the image of the head was moved while the background remained stationary, the response pattern indicates that the cell responded to the compatible movement of left view moving to the left, but not to other combinations of view and direction. For Figure 7.10,d-f the head remained stationary and the background was moved.

In condition 7.10d, the head was presented in the preferred view, and the background was moved in preferred direction of motion. Under these conditions the cell response was greatly reduced compared to the condition when the head itself was moving with preferred view and direction. This suggests that the cell was responsive to the head moving and not just the presence in the image of the appropriate form and the appropriate motion information. The response, however, was present when the background was moved in the opposite (null) direction behind the appropriate *static* head view (7.10e). This result indicates that the cell was sensitive to the relative motion of the contours of the head. Overall the result suggests that the cell was responding to the "motion energy" (Sugita & Tanaka 1991) of the object itself, and not simply a combination of form and motion information arising in the same retinal location.

Histological location of the recorded cells

The results of the histological reconstruction of one monkey (J) and the location of the cells conjointly selective for form and motion in the right hemisphere are shown in Figure 7.11. The figure presents a series of 1 mm sections running from 4.5 to 9.5 mm anterior to the inter-aural plane. The square symbols indicate the reconstructed positions of the cells selective for form and motion. The thick lines mark the cortical surface, the thin lines the inner edge of

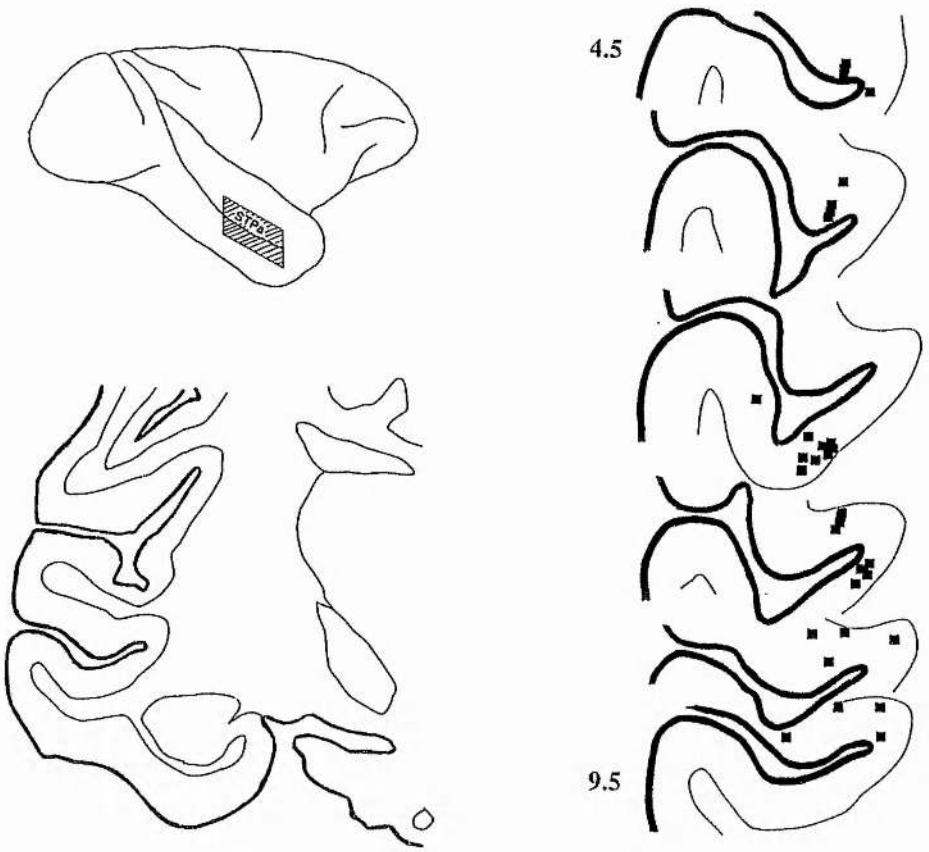


FIGURE 7.11. HISTOLOGICAL RECONSTRUCTION OF THE LOCATION OF CELLS SENSITIVE TO FORM AND MOTION. LEFT, UPPER: Schematic diagram of right view of the macaque brain. Shaded area shows location of area STPa. LEFT, LOWER: Section (9.5 mm anterior to the inter-aural plane) of the right hemisphere of monkey J. Shaded area indicates area STPa. RIGHT: Six 1 mm sections (4.5, 5.5, 6.5, 7.5, 8.5 and 9.5 mm anterior to the inter-aural plane) showing the superior temporal sulcus of the right hemisphere of monkey J are shown. The thick line indicates the cortical surface, the thin lines mark the edge of the grey matter. The square symbols indicate the location of recorded cells which showed selectivity to form and motion information.

grey matter. Cells showing sensitivity to form and direction were found in the upper bank, fundus and lower bank of the superior temporal sulcus. Cells from other monkeys were found in the same cortical regions.

DISCUSSION

Summary of the results

We have described here the response characteristics of cells in area STPa which show selectivity for combined form and motion information. The number of cells found suggests that this population of "form and motion" sensitive cells is as large as the population previously reported in area STPa selective for form (e.g. faces, hands: Bruce et al. 1981; Gross et al. 1972; Desimone et al. 1984; Perrett et al. 1982, 1984, 1992; Baylis et al. 1985; Young and Yamane 1992) and the population selective for direction of motion but not form (e.g. Bruce et al. 1981; Mistlin and Perrett 1990; Perrett et al. 1985b; Hietanen and Perrett 1993; Oram et al. 1993).

A small number of cells were found that showed limited selectivity to form and direction. These cells could be split into two groups. In both cases it should be noted, however, that the responses were dependent on motion and at least some form information. The responses of one of these populations showed partial selectivity for form since the responses of these cells was greater for bodies than for control objects. However the degree of form selectivity was limited, since there was no sensitivity to body view. The second population showing partial selectivity for form and direction showed clear form discrimination, preferring one body view to another, yet showed only partial motion selectivity, since the preferred body view moving in any direction produced equivalent responses.

The vast majority of the cells in area STPa showed more marked selectivity for both the form and the direction of motion of visual stimuli (Figures

7.4-7.9). Most of these cells were selective for only one view and direction combination, and therefore signal precise information about the object and its direction of motion relative to the viewer. Only a very small number of cells showed object-centered coding of view and direction. Importantly these cells did not supply a signal about "a moving body", but rather "a body moving forwards (compatibly)" or "a body moving backward (incompatibly)".

While not extensively tested, some cells exhibited evidence of 'binding' and required the correct relative motion of the form contours and not simply the presence of these contours combined with the preferred motion direction arising from other stimuli (i.e. Figure 7.4). Selectivity for motion direction defined by relative movement across a boundary has been described for area MST (Sugita and Tanaka 1991). These cells of MST, however, did not show selectivity for complex form that is described here for area STPa.

Coding of object motion and social signals

It is of interest to speculate on the possible functional roles of area STPa cells sensitive to form and motion. The most efficient single cell coding system to indicate what object is moving would be to have one cell population responding to the object, independent of the view-point, providing it was moving (i.e. body move in any direction). Although it is possible that "higher" visual areas may contain more cells with such selectivity, it is clear that at the level of area STPa such information is only available from combining the output of sub-populations since most cells were found to be selective for one combination of body view and direction. The fact that most cells code particular types of walking (compatible or incompatible but not both) does not mean that the general information that a body is present and moving is not evident from these neurons, rather that this information is only present when considering the combined activity of several cell populations. Note only 8 cell types are needed to fully represent body walking and

as such this is a sparse representation, with each element containing very specific information. This type of representation should not be likened to a distributed representation, where many units supply small amounts of information to a wide variety of tasks. The observation that nearly all STPa cells coding body view and direction were specific to particular combinations therefore suggests that they may serve a second, alternative, role.

An organism that is an active explorer of its habitat, such as the macaque, does by definition interact with the environment and other animals within it. Self movement in the environment requires optic flow information to be processed. Cell selectivities in areas MT and MST may well play an important role in this function (Duffy & Wurtz 1991a,b; Tanaka & Saito 1989). In many cases, however, information about position and speed of motion needs to be combined with identification of the moving object (predator/prey, subordinate/dominant troop member, possible mate or not etc.). Under these circumstances the manner of movement (compatible/incompatible) and the direction of motion are highly relevant.

Perhaps the most important situation for knowledge of view and direction of motion is during social interactions with conspecifics. Monkeys show a large and complex range of social behaviours, most of which involve body motions. Compatible movement away of a monkey from another can occur in many situations, but reversing away while still facing another individual indicates a more subordinate posture (Bertrand 1969). Thus signals of body view and direction of motion towards and away from the observing individual are important social cues. It has also been shown that monkeys can re-adjust their position within their troup's hierarchy on the basis of observed interaction between other individuals. For this to occur the monkey must be able to interpret interactions of others as a passive observer. Visual processing that leads to neural information about body view and direction of motion other than towards and away are

therefore also behaviourally useful, especially when combined with a "goal" of the direction of motion (Perrett et al. 1989).

Studies of cell selectivities in the amygdala and the effects of amygdalectomies suggest that this sub-cortical area is very important in generation of emotional and social responses (Brothers and Ring 1993; Aggleton 1993). Interestingly area STPa sends a strong projection to the amygdala (Aggleton et al. 1980; Turner et al. 1980; Amaral et al. 1992). Thus the cell selectivities reported here could provide visual information to the amygdala relevant to understand social signals.

It has been suggested that the role of cells in area STPa sensitive to particular views of the head (and particularly those sensitive to eye gaze) may also have a role in signalling the direction of another attention (Perrett et al. 1992, 1993). Here this idea is extended to include interpretation of body motion related cues in socially significant situations.

The existence of a large visually responsive cell population sensitive to both body form and direction of body motion in area STPa indicates that processing of visual information in the macaque, while for the most part occurring in two functionally distinct pathways, converges and establishes a conjoint representation of object identity and direction of motion. The response characteristics of this cell population can also be viewed as providing useful information for the interpretation of socially important signals, both from interactions with another monkey, and for the interpretation of behaviour between two other individuals. Such conjoint coding of two visual attributes might also underlie the unity in our every-day experience of perceiving objects *and their* motion.

CHAPTER 8

INTEGRATION OF FORM AND MOTION IN THE ANTERIOR SUPERIOR POLYSENSORY AREA (STPa) OF THE MACAQUE MONKEY II. POSSIBLE INTEGRATION METHODS

(Oram & Perrett, 1995b, *J. Neurophysiol.* in submission)

INTRODUCTION

Cells within the anterior section of the superior temporal polysensory area (STPa) show conjoint sensitivity to form (body view) and motion (direction) information (Bruce et al. 1981; Perrett et al. 1985a,b, 1989, 1990; Chapter 7). Moving body stimuli have classified into compatible (walking forwards where the body moves following one's nose) and incompatible movement (walking backwards). The majority of cells (125/161 cells, Chapter 7) responded to either compatible or incompatible body movement but not both. These cells all responded to one body view moving in one direction, but not to the same body view moving in a second direction, nor did they respond to a second body view moving in the preferred direction. The cells therefore showed clear response selectivity to conjoint visual form and direction information (see Chapter 7).

Sources of information

Chapter 6 documents cell selectivity in area STPa to form and motion where the form is defined purely by the motion of the stimulus ('biological motion' or 'moving light displays'. It is important to note, however, that only 25% of cells showed selectivity to motion defined form, and that the majority of these

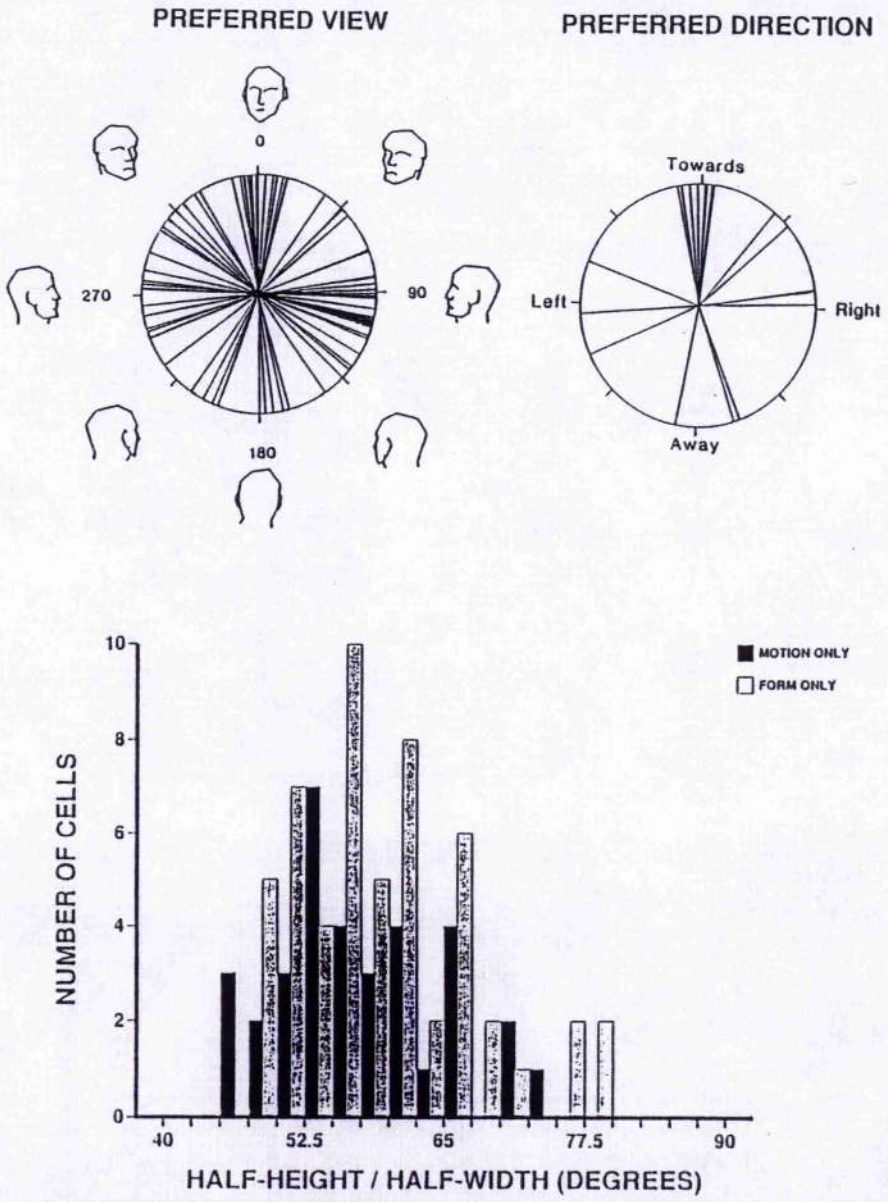


FIGURE 8.1. SIMILARITY OF RESPONSE CHARACTERISTICS OF "FORM ONLY" AND "MOTION ONLY" CELLS. **UPPER: SIMILARITY OF PREFERRED VIEWS AND DIRECTIONS.** The left section shows a polar plot of preferred body views for 73 cells selective for the static view of the head and body. The right section shows a similar plot of the preferred direction of cells selective for object motion in the horizontal plane independent of the object form. In both cases there is significant clustering around the cartesian axes (Perrett et al. 1991; Oram et al. 1993). **LOWER: SIMILARITY OF WIDTH OF TUNING.** The dark bars show the half-height/half-width measurement of tuning curves fitted to data collected from "motion only" cells (Oram et al. 1993). The light bars show the same measurement from "form only" cells (Perrett et al. 1991). These two populations are statistically indistinguishable (Mann-Whitney U, $p = 0.24$).

(78%) showed reduced response magnitude compared to the responses obtained under normal light conditions (Chapter 6). Further, those cells which showed selectivity to biological motion also responded to translation of the body (Perrett et al. 1990a,b). Indeed, of cells selective for walking stimuli which were tested with translating bodies all (13/13, Harries, Perrett and Oram unpublished studies) maintained form selectivity (body view) to stimuli translating in the preferred direction. Under these conditions the relative motion inputs of the articulating limbs cannot be used to establish form selectivity (see Chapter 6 for discussion). Therefore, although some 25% of cells sensitive to form and motion can perform form-from-motion, these same cells also integrate information from both form signals and motion signals. Further, only 6% of cells tested with biological motion stimuli responded as strongly as when tested under normal lighting. This suggests that the degree of the role of form-from-motion processing, while present, is relatively small compared to the integration of form signals with motion signals.

Reviewing published material concerning other types of neural populations in area STPa indicates ways from and motion signals could be combined. Figure 8.1 shows two ways in which the response characteristics of two other cell populations in area STPa would facilitate the integration of those two sources of information. As can be seen in the upper section of Figure 8.1, cells sensitive to direction of motion but not form show a tendency to be maximally responsive to directions of motion along the cardinal axes (towards away left and right, Chapter 5). Cells tuned for different static views of the head and body also show significant clustering about particular views. In particular the face, left and right profiles and back views of the head are more strongly represented than the intermediate views (Perrett et al. 1991). Signals from these cells could thus be combined to establish cells with conjoint selectivity for form and motion in such a way as to be responsive to either compatible or incompatible walking motion.

The lower half of Figure 8.1 shows a second way in which the response characteristics of cells selective for form but not motion and cells selective for direction of motion independent of form could facilitate their integration. The half-height half-width measure for these two cell populations are very similar. The implications for this are best illustrated by using an example. Imagine a sub-population of cells all sensitive for motion of any object towards the observer, and a second sub-population of cells sensitive to the front view of the head and body. When viewing an approaching body, both cell sub-populations would be firing maximally. When viewing a body walking towards and slightly to the left, both cell sub-populations would have reduced firing rates. Importantly though, the reduction in firing rates would be equivalent. Thus the strength of signal between these populations would co-vary in an equivalent way.

Linearity of information integration

Cells showing conjoint sensitivity to form and motion appear to bind these two types of information such that the motion signal has to derive from the object itself (Chapter 7). While there is no attempt to investigate directly the phenomenon of binding in this study, some aspects of the integration are examined. One method for generating form and motion selectivity would be to simply sum an input response to object movement in the appropriate direction with an input response to the appropriate body form. Such summation could either be linear or show non-linearities. Under non-linear summation, the response to these two components (motion direction and view) when presented simultaneously would not equal the arithmetic sum of the response magnitudes to these two components presented in isolation.

It is clear that neural processing in many visually responsive cortical areas is highly non-linear. Non-linearity is evident in striate cortex where a non-linear relationship exists between contrast and response (contrast gain control, for review see Albrecht and Geisler 1994). Perhaps even more obvious is the non-

linearity described for many visually responsive cells to stimulation outside the classic receptive field (e.g. Allman et al. 1985; Knierim and Van Essen 1992). Positional invariance is seen in the cell responses found in inferotemporal cortex (Gross 1992; Tovee and Rolls 1993), again indicating a non-linear system. From theoretical considerations it is also desirable to have non-linear mechanisms. This is because many operations, such as positional invariance, are not possible using single layer networks (Minsky and Papert 1969). As all linear networks can be described using an equivalent one layer network, this implies that linear integration could not produce such operations.

Evidence of integration of form and motion could be associated with one of four response patterns: (1) In the simplest situation the cell would show little or no response to either the correct form (presented statically) or to the preferred direction of motion (independent of form), but would show a strong response to the conjoint information, (2) the form signal could produce a strong response (approximately equal to the response to the conjoint presentation), with little or no response to the motion signal, but response suppression when the motion signal was in the non-preferred direction, (3) the motion signal could produce a strong response (approximately equal to the response to the conjoint presentation), with little or no response to the static preferred form signal, but response suppression when there was a non-preferred form signal, and finally (4) both the form signal and the motion signal could produce large responses when presented in isolation. In the first case, the non-linearity is due to response enhancement, while in situations 2-4, any non-linearity that might be present is due to response suppression.

In this chapter, possible mechanisms by which integrated form and motion selectivity could be established are considered. In particular evidence for non-linearity in the integration process is examined. Response characteristics of cells showing conjoint sensitivity to form and motion with cell populations within area

STPa that are sensitive to (a) form but not motion and (b) motion but not form are also compared.

METHODS

The methods used for data collection have been given in detail in the general methods and the previous chapter.

The additional analysis procedures adopted are described below. Three indices were calculated for each cell. These were a Motion sensitivity index (I_m), a Form index (I_f) and a Linearity index (I_l). I_m indicates the ability of a cell to discriminate between its preferred view/direction combination and the static image of the preferred view (i.e. I_m gives a measure of the effect of stimulus motion on the cell response). For each cell I_m was calculated as the difference in response magnitude between the preferred view/direction combination and the static presentation of the preferred view. To allow comparison between cells, this difference was normalised by dividing it by the response magnitude to the preferred view/direction combination. The Motion index (I_m) is

$$I_m = \frac{(\text{Pref} - \text{SA}) - (\text{Static} - \text{SA})}{(\text{Pref} - \text{SA})}$$

where SA is spontaneous activity, Pref is the measured response to the preferred view/direction combination, and Static is the measured response to the static presentation of the preferred view.

A value of 0 for I_m indicates that the response magnitude to the moving and static stimuli are equivalent. When $I_m = 1$, this indicates that there was no response to the preferred view when static. A value of 0.8 indicates that 80% of

the response to the preferred view/direction stimulus is *not* explained by the response to the static preferred view. Values less than 0 indicate that the response to the preferred view/direction combination was less than the response to the static preferred view. Values of I_m which are greater than 1 indicate inhibition relative to SA for the presentation of the static preferred view.

In a similar manner, the Form sensitivity index (I_f) was calculated as the normalised difference between the response magnitudes to the preferred view/direction combination and control objects moving in the preferred direction, giving

$$I_f = \frac{(\text{Pref} - \text{SA}) - (\text{Control} - \text{SA})}{(\text{Pref} - \text{SA})}$$

where Control = measured response to control objects moving in the preferred direction, other definitions as for I_m).

Calculation of the linearity index (I_l) was performed using

$$I_l = \frac{(\text{Pref} - \text{SA}) - ((\text{Static} - \text{SA}) + (\text{Control} - \text{SA}))}{(\text{Pref} - \text{SA})}$$

(definitions as for I_m and I_f). Here it is the difference between the response magnitude to the preferred view/direction combination and the (linear) sum of the responses to the static preferred view and to controls moving in the preferred direction that has been normalised. This means a value of 0 indicates that linear summation of the form response and direction response can explain the response to the preferred view/direction combination. When $I_l = 1$, then the sum of the responses to the static preferred form and control moving in the preferred direction is no different from spontaneous activity and the cell is summing inputs

in a highly non-linear way (with respect to static view and control direction of motion responses).

To allow comparison with cells selective for stimulus form but not selective for motion, a view discrimination index (I_V) was calculated.

$$I_V = \frac{\text{(Pref - SA) - (Opp View - SA)}}{\text{(Pref - SA)}}$$

where Opp View = the measured response to the body view opposite to that defined in the preferred stimulus moving in the same preferred direction. Note that this index compares responses where the only stimulus parameter that has changed is the body view.

The direction discrimination index (I_d) was calculated as

$$I_d = \frac{\text{(Pref - SA) - (Opp Direction - SA)}}{\text{(Pref - SA)}}$$

where Opp Direction = the measured response to the preferred body view moving in the direction opposite to the preferred direction. Again note that this index compares responses where the only stimulus parameter that has changed is the direction of motion.

RESULTS

Of the 125 cells found to show unimodal sensitivity to body view (form) and direction of motion from the four subjects, examples of all four methods of

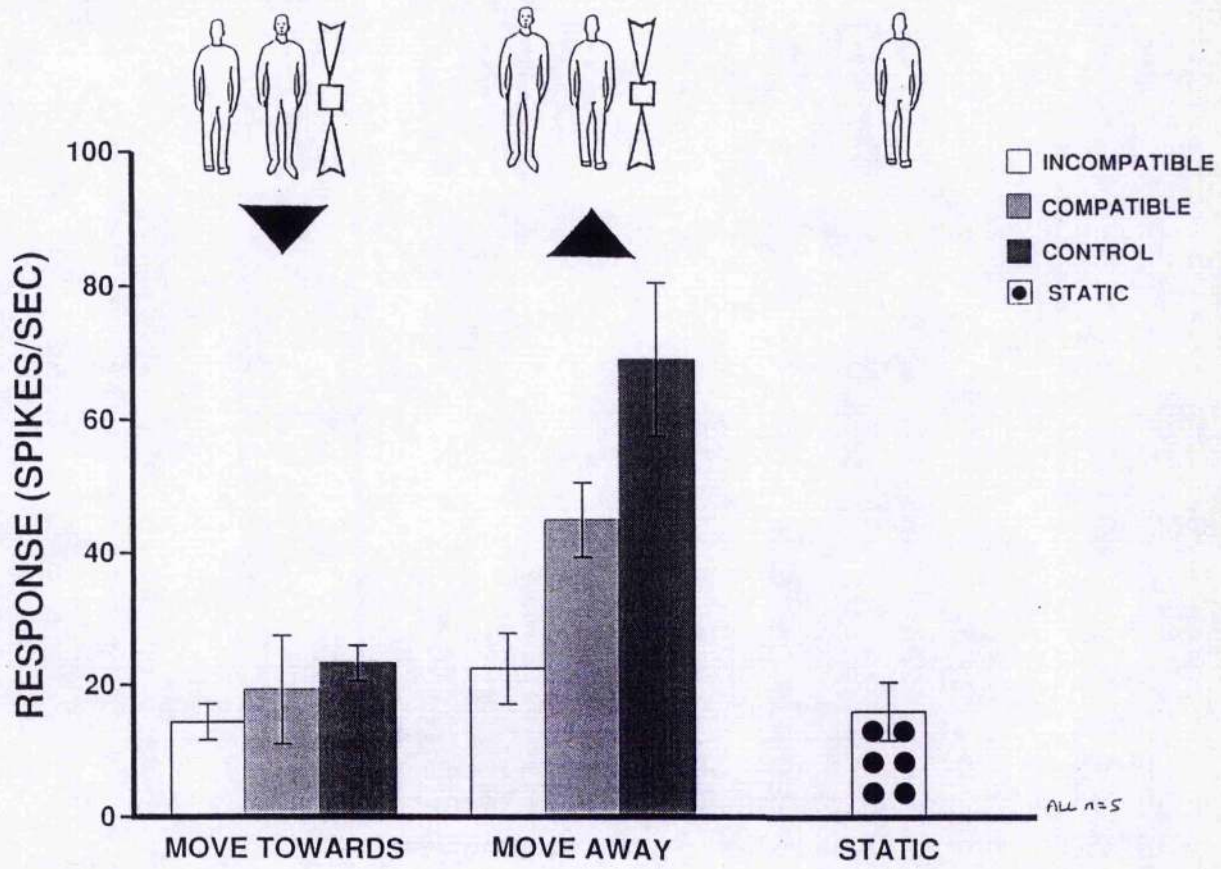


FIGURE 8.2. CONJOINT FORM AND MOTION SELECTIVITY: RESPONSE SUPPRESSION TO THE INAPPROPRIATE BODY VIEW. Mean response magnitude (+/- S.E.M.) are shown for the one cell which showed response suppression to the inappropriate form. The stimuli are shown schematically above the histogram bars. The cell responded well to control object movement away from the monkey. When the back view of the body moved away the response was reduced compared to control level ($p = 0.002$) and was reduced further when the front view of the body moved away ($p = 0.003$). The response to the static back view was no different to spontaneous activity (not shown). [Overall effects of 2-way ANOVA on moving stimuli: Direction of motion $F_{[1,24]} = 23.1$, $p < 0.0005$; Type of motion $F_{[2,24]} = 8.4$, $p = 0.002$; Interaction $F_{[2,24]} = 3.9$, $p = 0.034$.]

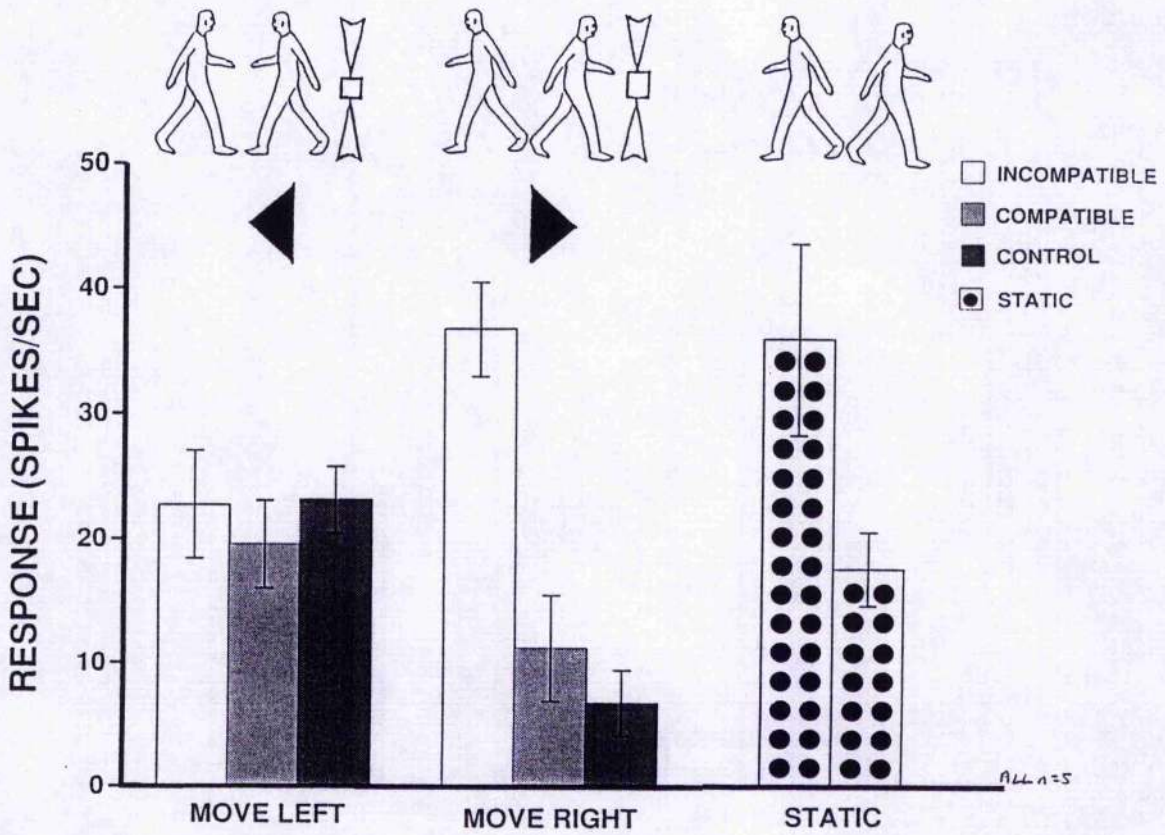


FIGURE 8.3. CONJOINT FORM AND MOTION SELECTIVITY: RESPONSE SUPPRESSION TO THE INAPPROPRIATE DIRECTION OF MOTION. Mean response magnitude (+/- S.E.M.) are shown from a single cell which showed response suppression to the inappropriate direction of motion. The stimuli are shown schematically above the histogram bars. The cell responded well to images of the left profile either when static or moving to the right (or towards or away, not shown). When the left profile was moved to the left the response was reduced ($p < 0.0005$). [Overall effect of conditions: View $F_{[3,64]} = 1.9$, $p = 0.138$; Type of motion $F_{[3,64]} = 8.6$, $p < 0.0005$; Interaction $F_{[9,64]} = 4.0$, $p < 0.0005$.]

integration postulated in the introduction were present. The relative numbers of each were, however, very different. Only one cell (1/125, <1%) was classified as showing response suppression when the inappropriate body view moved in the appropriate direction. This cell is shown in Figure 8.2. This cell showed a unimodal response pattern for compatible movement away from the subject, since the response to the back body view moving away was greater than the responses elicited by either the front view moving away from or the back view moving towards the subject. As can be seen, the cell produced the strongest response when control objects were moved away from the subject. Differential responses were seen to different views of the body, however, when the body walked away from the monkey. Slight suppression was seen when the back view of the body moved away compared with control object motion. The degree of suppression was significantly greater when the front view of the body moved away (incompatible walking). For this one cell it is clear that object movement away produced a large response unless the stimulus "object" was the front view of the body, in which case the cell did not respond. Comparison with the response observed to the static image of the back of the body (and front view, not shown) indicates that this effect is indeed due to suppression or prevention of the response to the direction away signal.

A further five cells (5/125, 4%) were found that showed response suppression to an inappropriate direction of motion. An example of this type of response pattern is shown in Figure 8.3. This cell was classified as showing unimodal selectivity for incompatible movement to the right (i.e. left profile moving right). For this cell it is clear that the response pattern showed view selectivity for static images of the body. It is clear that the unimodal selectivity for moving bodies is due to response suppression when the preferred view (left profile) moves in the non-preferred direction (to the left).

The remaining cells (119/125, 95%) showed either linear summation of responses to form and motion inputs or response enhancement. Figure 8.4 shows

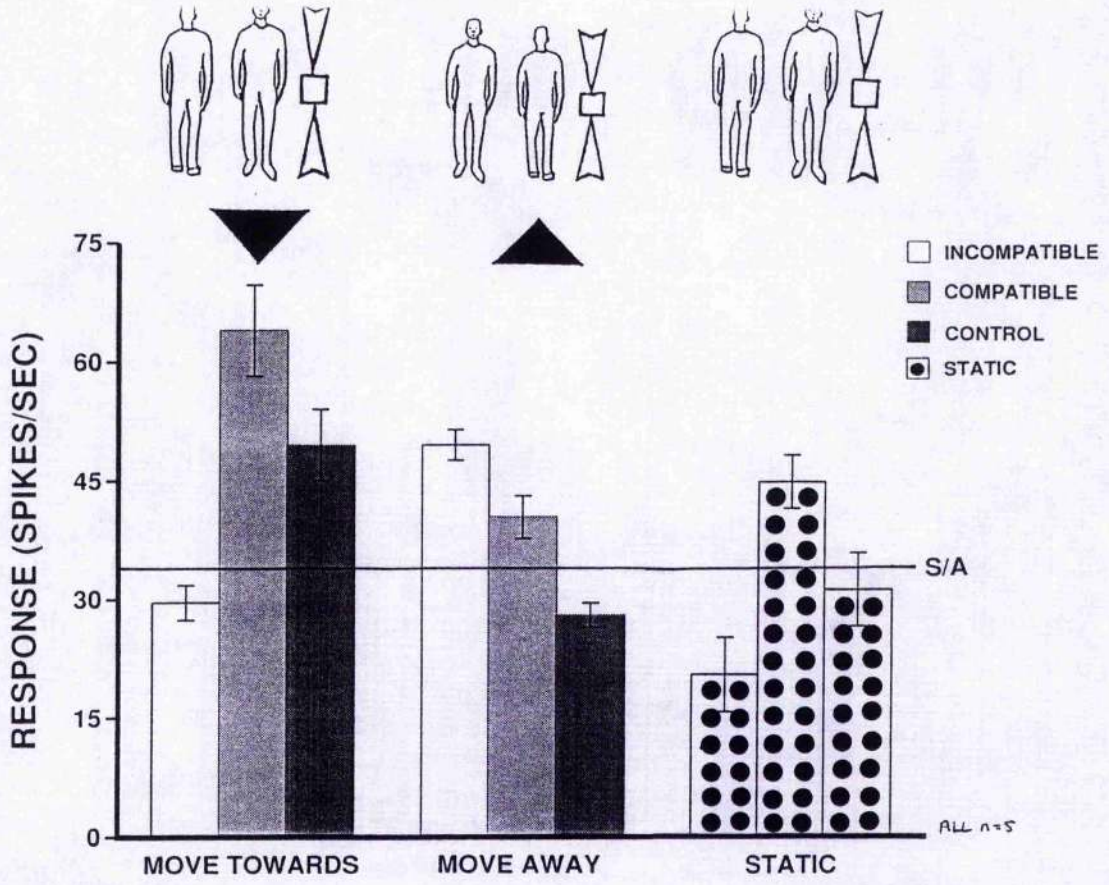


FIGURE 8.4. CONJOINT FORM AND MOTION SELECTIVITY: LINEAR RESPONSE ADDITION. Mean response magnitude (+/- S.E.M.) are shown for one cell which showed linear response summation. The stimuli are shown schematically above the histogram bars. S/A, spontaneous activity. The cell responded better to images of the front view when moving towards the monkey than when moving away or static ($p < 0.0005$ each comparison). Taking response magnitudes above S/A, the cell showed linear response summation for movement towards the monkey: The response to the front view of the body walking towards the monkey is approximately equal to the sum of the responses to the static front view of the body and control objects moving towards the monkey (30.0 approx= 10.8 + 15.6). Similalry the response to the back view moving towards the monkey is approximately equal to the sum of control object motion towards the monkey and the response to the static back view of the body (2.0 approx= 15.6 + -13.6).

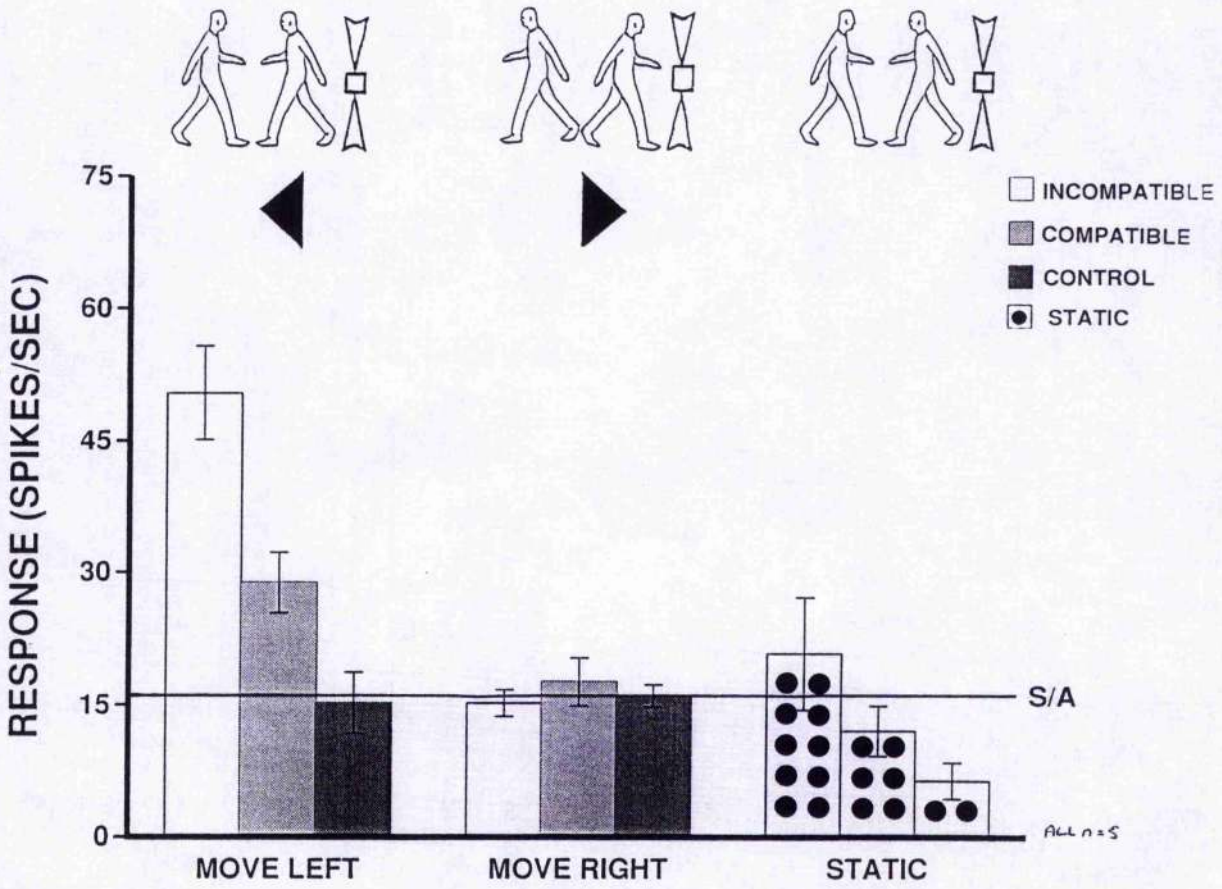


FIGURE 8.5. CONJOINT FORM AND MOTION SELECTIVITY: NON-LINEAR RESPONSE ADDITION. Mean response magnitude (+/- S.E.M.) are shown for one cell which showed approximately the average non-linearity in response summation. The stimuli are shown schematically above the histogram bars. S/A, spontaneous activity. The cell responded better to images of the right profile moving to the right than other body view or control direction combinations (compatible and incompatible movement in 4 directions were tested, $p < 0.0005$ each comparison; Overall effects Direction $F_{[3,48]} = 25.0$, $p < 0.0005$; Type of motion $F_{[2,48]} = 3.5$, $p = 0.04$; Interaction $F_{[6,46]} = 10.9$, $p < 0.0005$).

an example of a cell with linear summation of responses. This cell showed a unimodal response selectivity for compatible movement towards the subject. Taking spontaneous activity (S/A) as the baseline activity level, it is clear that arithmetic summation of the response to the static presentation of the front body view (10.8 spikes/sec) and the response to control objects moving towards the subject (15.6 spikes/sec) is approximately equal to the response obtained when the front body view walks towards the subject (26.4 approx= 30.0 spikes/sec). Furthermore, the response to control objects moving towards the subject (15.6 spikes/sec) plus the (inhibitory) response to the static view of the back of the body (13.6 spikes/sec below S/A) is approximately equal to the response seen when the stimulus is the back of the body walking towards the subject (2.0 approx= -4.4 spikes/sec).

The most common response pattern indicated non-linear summation, with response enhancement occurring when the preferred view was combined with the preferred direction of motion. Figure 8.5 gives a typical example of this type of cell. This cell showed a strong response (34.4 spikes/sec above S/A) to incompatible movement to the left (right profile walk left). Control movement to the left was 0.8 spikes/sec below S/A, whereas the response to the static view of the right profile was 4.8 spikes/sec above S/A. For this cell therefore the response to the preferred view and direction combination was over eight times the sum of the responses to the separate view and direction components (34.4 vs 4 spikes/sec).

Sensitivity Indices

The response to the static image of the preferred body view was measured for 52 cells. Figure 8.6 shows the distribution of the Motion sensitivity index (I_m, see methods) for these cells. Ignoring the five cells which showed response suppression when the preferred view was moved in the null direction (solid bars), the range of I_m is 0.28 to 1.98 (mean 0.83, S.D. = 0.31, median = 0.77). For this

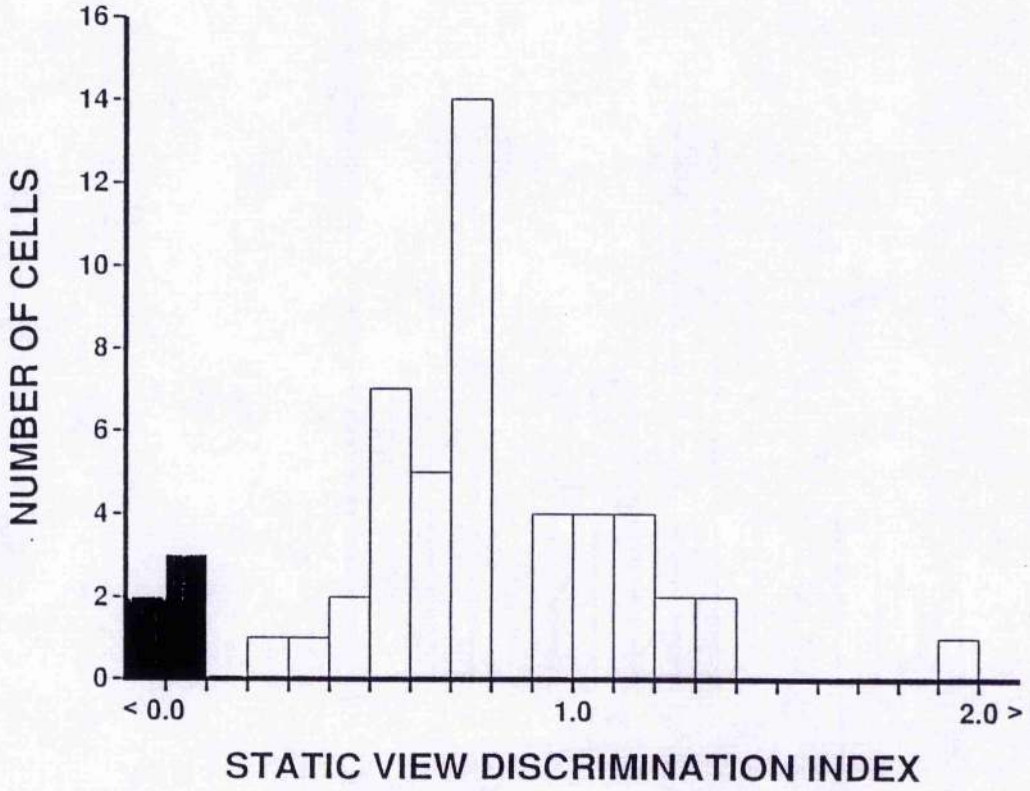


FIGURE 8.6. DISTRIBUTION OF THE STATIC VIEW DISCRIMINATION INDEX. The static view discrimination index (I_S) was calculated as 1 - (Static - SA)/(Pref - SA), where SA is spontaneous activity, Pref is the measured response to the preferred view/direction combination, and Static is the measured response to the static presentation of the preferred view. The dark bars show those cells which showed response suppression when the image of the preferred body view was moved in the non-preferred direction. See text for details.

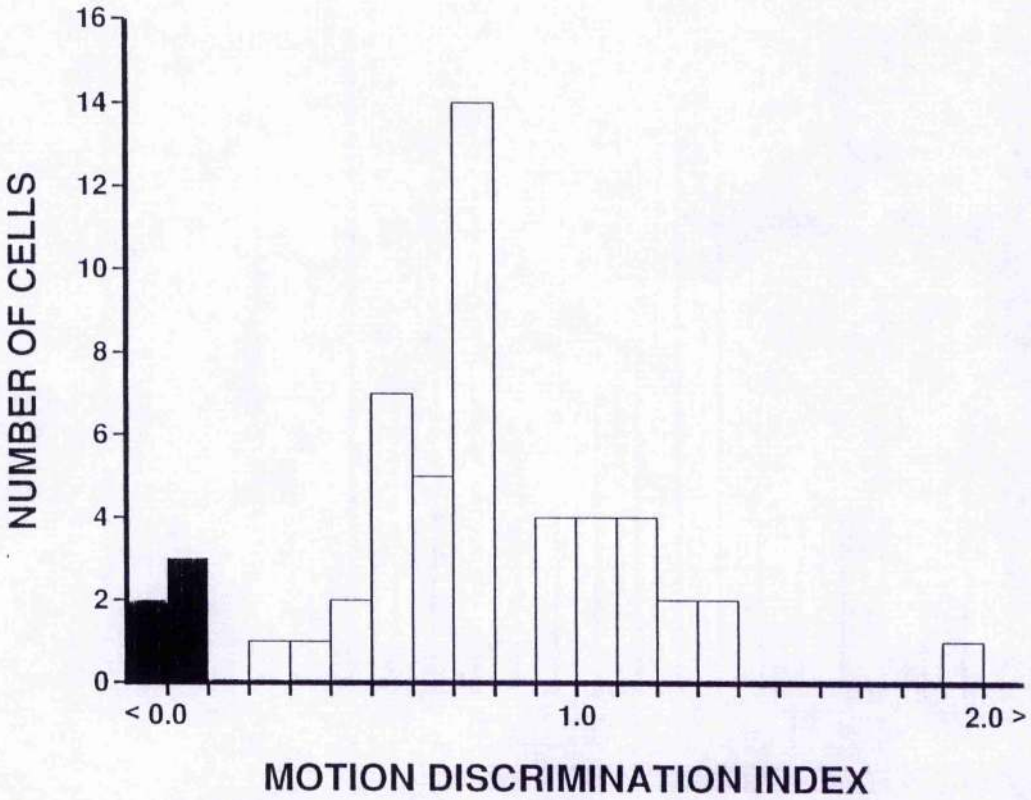


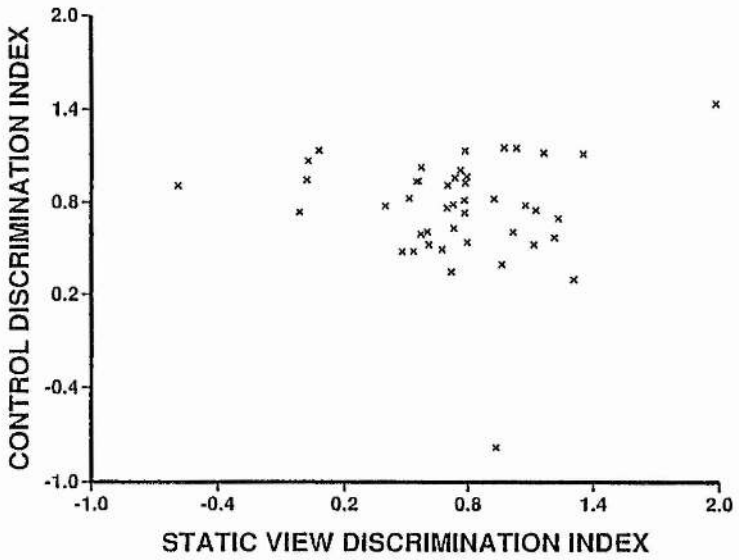
FIGURE 8.7. DISTRIBUTION OF THE MOTION DISCRIMINATION INDEX. The control discrimination index (I_m) was calculated as 1 - (Control - SA)/(Pref - SA), where SA is spontaneous activity, Pref is the measured response to the preferred view/direction combination, and Control is the measured response to control objects moving in the preferred direction. The dark bar shows the cell which showed response suppression when the image of the non-preferred body view was moved in the preferred direction. See text for details.

latter population of cells, discrimination between the preferred body view static and moving in the preferred direction is high, with only 17% of the response to the preferred view/direction combination being accounted for by the response to the static presentation. Alternatively this index can be regarded as giving a measure of the impact of motion in the preferred direction: adding motion to the preferred stimulus form gives a response that is, on average, 1/0.17 = 5.9 times the response to the same stimulus without motion.

Response magnitudes to control objects moving in the preferred direction were measured for 83 cells. The distribution of the Form sensitivity index (I_f) are shown in Figure 8.7. When the one cell that responded strongly to control objects moving in the preferred direction (solid bar, see also Figure 8.2) is ignored, the range of I_f is 0.21 to 1.44 (mean 0.80, S.D. = 0.25, median 0.79). For this sample of 82 cells, only 20% of the response to the preferred view/direction combination can be accounted for by motion input alone. The effect, therefore, of adding appropriate form information to a stimulus moving in the preferred direction, on average, to increase the response magnitude five-fold.

Independence of form and motion information

To investigate the possibility that there was some level of interdependence of the form and direction sensitivity in the responses, the values of I_m and I_f were correlated. Figure 8.8 upper shows the scatter plot of the 44 unimodal cells for which both indices were calculated. As is clear from Figure 8.8, there is no systematic relationship between I_m and I_f ($r_{[42]} = -0.02$, $p = 0.9$). This lack of correlation was also found when the values corresponding to those cells which showed response suppression, either by the inappropriate direction (n=5) or by the inappropriate form (n=1) were removed ($r_{[36]} = 0.293$, $p = 0.07$). Although this nearly reaches significance, much of the trend in the correlation can be attributed to one "out-lier" ($I_m = 1.98$, $I_f = 1.44$). When this value is also removed, the correlation drops markedly ($r_{[35]} = 0.056$, $p = 0.74$, Figure 8.8 lower).



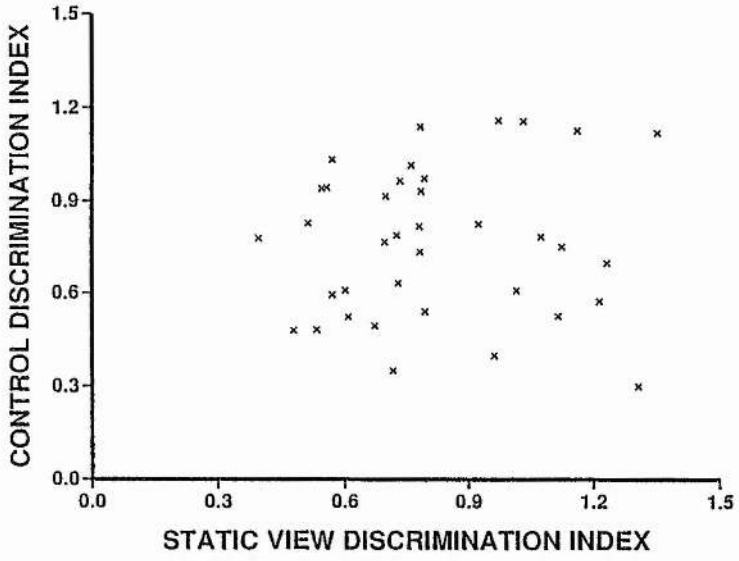


FIGURE 8.8. CORRELATION OF THE MOTION DISCRIMINATION AND THE STATIC VIEW DISCRIMINATION INDICES. UPPER: The calculated values of I_m ans I_f are shown for all cells where data was available. The lower figure shows the plot of I_m and I_f for a subset of these cells. See text for details.

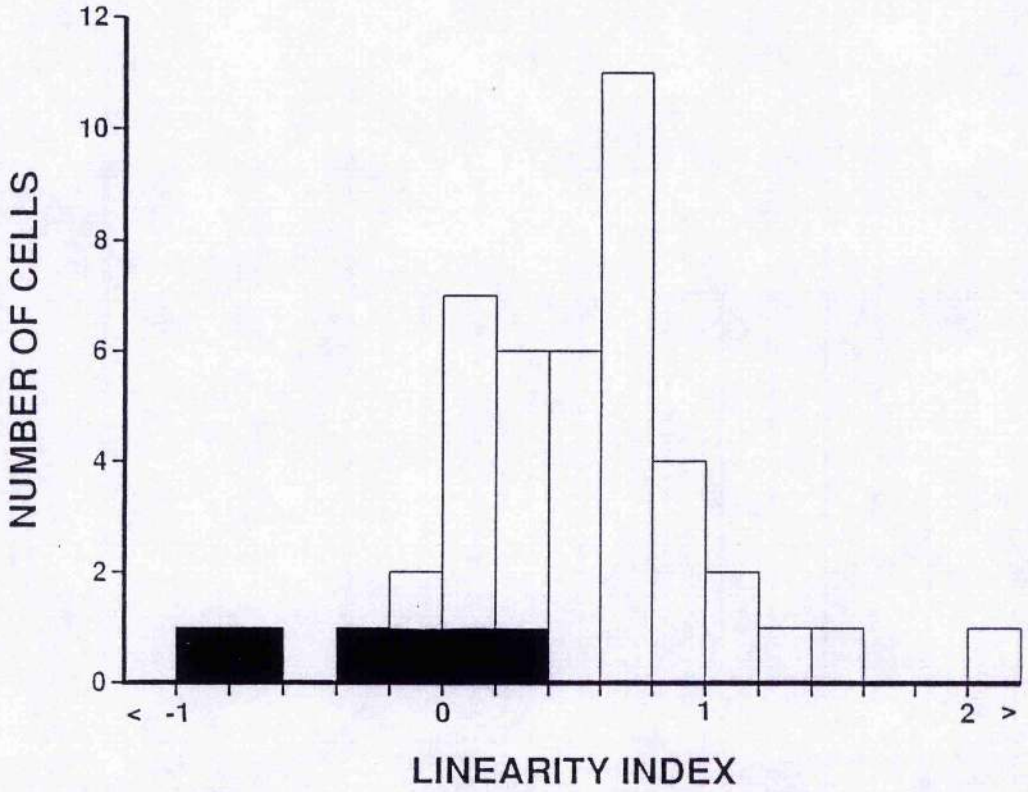


FIGURE 8.9. DISTRIBUTION OF THE LINEARITY INDEX. The calculated values of I_l are shown. The linearity index, (I_l) was calculated as 1 - {(Control - SA)+(Static - SA)}/(Pref - SA), where SA is spontaneous activity, Pref is the measured response to the preferred view/direction combination, Control is the measured response to control objects moving in the preferred direction and Static is the measured response to the static presentation of the preferred view. The dark bar shows the value of I_l of those cells which showed response suppression either to the inappropriate body view or direction or motion. See text for details.

Linearity index

The linearity index (I_l) was computed for 44 cells. For the other cells an insufficient number of conditions were tested. The distribution of I_l is shown in Figure 8.9. I_l ranged from -0.84 to 2.4 (mean 0.51, S.D. = 0.55, median 0.52). Removal of those cells which showed response suppression (solid bars), either to the inappropriate body view or direction or motion, gives the range if I_l is 0 to 2.4 (mean 0.63, S.D. = 0.46, median 0.60). From this sample estimate it appears as if the average response enhancement from the conjoint form and motion information is some 2.7 times the response to the sum of the isolated form information and direction information signals. Nearly one half (16/38, 42%) of the cells showed at least a trebling of response magnitude compared to that predicted by linear summation.

Comparison with other STPa cell populations

It was noted in the introduction that two other cell populations exist in area STPa. One of these is sensitive to static form information, the second is sensitive to direction of motion but shows no selectivity for form. In this section the response characteristics of cells sensitive to conjoint form and motion information are further compared.

Cell response latencies in each of the three populations were assessed as the first of three consecutive bins that exceeded the 95% confidence interval of the pre-stimulus sample period (see Chapter 3 and 5). The mean response latencies (see Figure 8.10) of these three cell populations were found to vary significantly ($F_{[2,94]} = 7.35$, $p = 0.001$). Post-hoc testing revealed that cells which responded to motion (with or without conjoint form selectivity) had a shorter average response latency (99 and 90 ms) than the cells sensitive to static form (119 ms, $p < 0.02$ each comparison).

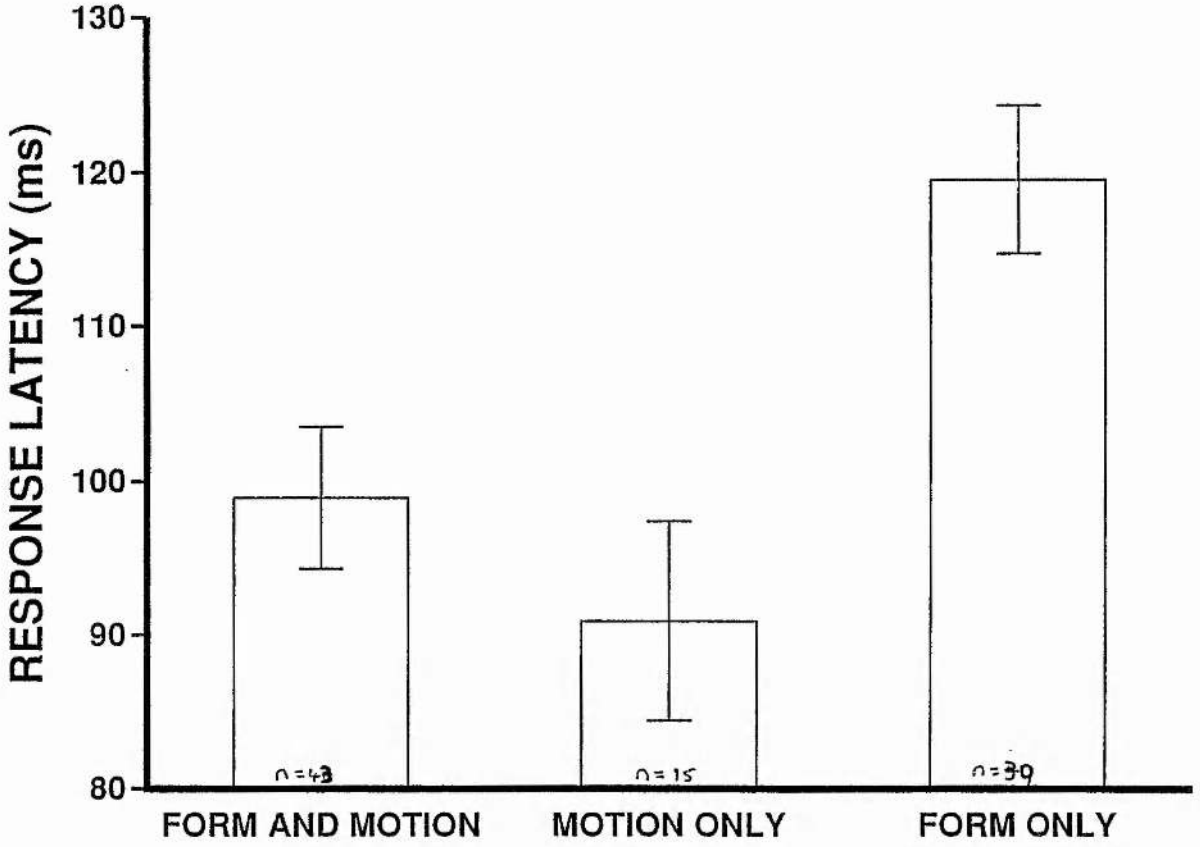


FIGURE 8.10. COMPARISON OF RESPONSE LATENCIES FOR CELLS SENSITIVE TO "FORM AND MOTION", "MOTION ONLY" OR "FORM ONLY". Response latency (+/- 1 S.E.M.) was assessed for cells selective for motion only (from Oram et al. 1993), form only (from Oram and Perrett 1992) and form and motion. Comparison of response latencies revealed that on average the responses of "form only" cells had a longer latency than "motion only" or "form and motion" cells ($p < 0.02$ each comparison). The average response latencies of "motion only" and "form and motion" cells were the same ($p > 0.3$). [Overall effect of conditions: $F_{[2,94]} = 7.35$, $p = 0.001$]

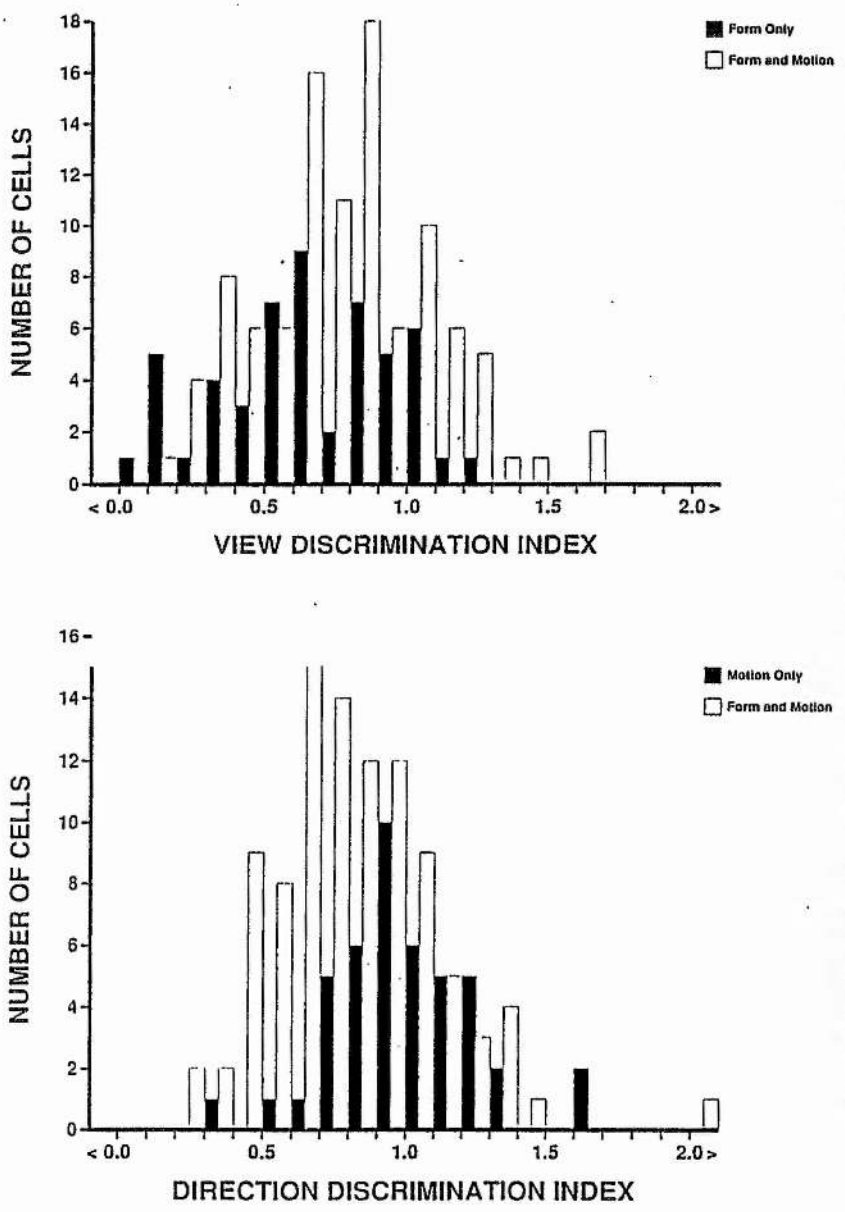


FIGURE 8.11. UPPER: VIEW (FORM) DISCRIMINATION INDICES FOR "FORM ONLY" AND "FORM AND MOTION" CELLS. The view discrimination index was calculated as 1 - (Opp_V-SA)/(Pref-SA), where SA = spontaneous activity, Pref is the response to the preferred view and Opp_V is the response to the view 180 degrees away from the preferred view. [For "form and motion" cells (dark bars) both stimuli were moving in the preferred direction.] These two discrimination index populations differ slightly ($p = 0.025$). LOWER: DIRECTION DISCRIMINATION INDICES FOR "MOTION ONLY" AND "FORM AND MOTION" CELLS. The direction discrimination index was calculated in an analogous way to the view discrimination index: 1 - (Opp_d-SA)/(Pref-SA) where Pref = response to the preferred direction of motion and Opp_d is the response to motion 180 degrees away from the preferred direction. [For "form and motion" cells (dark bars) the preferred view was used in both stimuli.] These two populations differ ($p = 0.001$).

The distribution of the view discrimination index (I_v) is shown in the upper section of Figure 8.11. The distribution of the equivalent index for cells sensitive to static form (Perrett et al. 1991) is plotted for comparison. Statistical comparison of these two populations indicates a small but significant ($t_{[151]}$ = 2.269, p = 0.025) difference between these populations, with the cells sensitive to conjoint form and motion showing slightly improved view discriminability (mean I_v = 0.78 +/- 0.031) than that shown by cells responsive only to static form (mean I_v = 0.66 +/- 0.041).

The lower section of Figure 8.11 shows the distribution of the direction index (I_d). The distribution of I_d for cells sensitive to motion but showing no form selectivity (Chapter 5) is also plotted. These two populations differ, with the cells sensitive to conjoint form and motion showing significantly ($t_{[115]}$ = 3.43, p = 0.001) poorer directional discrimination (mean I_d = 0.83 +/- 0.036) than that shown by cells responsive to motion independent of form (mean I_d = 1.01 +/- 0.037).

DISCUSSION

Response suppression vs response enhancement

The vast majority of cells (119/125, 95%) showed response addition or response enhancement. Very few cells (6/125, 5%) were found whose response selectivity indicated suppression when either a non-preferred form (Figure 8.2) or a non-preferred direction (Figure 8.3) was added to an adequate stimulus. There are several points to be made with respect to this issue. It is possible that the number of cells showing suppressive influences have been underestimated. The method of initial classification, which guided subsequent testing, was such that

cells of these types could have been wrongly classified as showing (a) no form sensitivity since they responded well to control object motion (Chapter 5), or (b) no motion sensitivity since they responded well to static images of the body (Perrett et al. 1991). The number of cells that may have been mis-classified in this way, however, is likely to be small, and certainly well below the number needed to make this type of information integration appear as common as response addition or enhancement (some 130 further cells). The initial testing procedures typically involve walking, both compatibly and incompatibly, in a variety of directions before the use of control objects. In doing this it is assumed that if a cell shows no form selectivity, then it will also respond to the walking body. It is only after observing responses to a walking body that control objects are tested. Thus differential responses to different body views moving in a particular direction was likely to be observed before control objects were tested.

Sensitivity indices

It is clear that the responses of this population of cells covers "walking-body" stimulus space in an efficient (sparse) manner with only eight cell sub-populations necessary (Chapter 7). Since competitive interactions between elements will lead to a sparse representation over stimulus space (Foldiak 1990) it is reasonable to assume that interactions between these representations will be more pronounced than competitive interactions between neural representations for walking bodies and other moving objects. Given that this assumption holds, the impact of motion on these cell responses is better investigated by comparing responses to static images of the preferred form (body view) rather than comparison with the same body view moving in other directions. Likewise, the use of control objects moving in the preferred direction allowed comparison of the effect of body view on the response while minimising possible interactions between cells selective for different body views moving in the same direction.

The values obtained for the Motion sensitivity index I_m indicated that the effect of adding motion in the preferred direction increased response magnitudes on average by nearly 6 times (mean $I_s = 0.83$). This shows a clear requirement of these cells for motion before a strong response is elicited. A similar dependency was found for form information, with the addition of preferred form increasing response magnitude to moving stimuli five fold (mean $I_f = 0.80$).

Independent form and motion signals

It is also relevant to consider the degree of separation between the "form" and "motion" pathways. The inferotemporal cortex (IT) classically belongs in the "form" pathway, yet responses in IT do show some sensitivity to motion. Sensitivity to shape defined by motion has been observed in cells of the inferotemporal cortex (Sary et al. 1993), yet these cells were not reported as showing directional preferences. This being the case, form inputs driven by motion of the shape against a background derived from IT cortex would not enable the directional selectivity of the STPa cells reported here. Evidence from a human neuropsychological case study indicates that some form sensitivity is perhaps directly available in the "motion" pathway (also called the "how" pathway, Goodale and Milner 1993; Milner and Goodale 1994). Recent studies have shown, however, that this form sensitivity is limited to orientation, size and "center of mass" information (Goodale et al. 1994).

In the previous chapter it was shown that the responses of the cells described here were dependent on conjoint form and motion information and not simply the close proximity of these two visual attributes. The evidence from the discrimination indices presented here shows that the absence of either one of these two signals leads to a marked reduction in response magnitude. These data do not, however, indicate whether or not there is any inter-dependency of these types of information. Studies of IT cortex showed many cells were selective for a combination of colour, shape and texture (Tanaka et al. 1991; Komatsu and Ideura

1993). While the IT neural responses were dependent on each of these attributes, there was no co-dependency of colour sensitivity with shape for the cells studied. Furthermore, the tuning curve for colour showed the same pattern, regardless of the shape in which the colour was presented (Komatsu & Ideura 1993).

If there was a dependency of form information on direction of motion, or conversely, a dependency of motion information on form, then it would be expected that there would be a relationship between the Motion and Form sensitivity indices. This was found to not be the case (Figure 8.8). A tentative conclusion that the sources of form information and motion information are independent can therefore be reached from the data presented here.

The degree of non-linearity

The investigation into the degree of linearity of the responses of these cells showed that the mechanism for combining form and motion information within area STPa is highly non-linear. The linearity of neuronal integration of two or more visual attributes has not been extensively studied where the incoming information is thought to have been processed largely independently of each other before being combined.

While for a few cells, the responses may be a simple linear summation of form and motion inputs, generally this is not the case. The proportion of the response to the preferred view and direction combination that can be related to the motion signals alone is on average 20%. A similar figure (17%) can be attributed to form information alone. It therefore appears as if the relative strength of inputs from these two sources is approximately equal.

Comparison with other STPa cell populations

The properties of two other STPa cell populations tuned to form alone or motion alone are such that these cells would seem to be suited to integration of form and motion (Figure 8.1). Comparison of the response latencies of the three

cell populations (cells showing selectivity for conjoint form and motion, static form or direction of motion independent of form) indicated a difference, with cells selective for static form having a longer response latency than the other two populations. This difference in latencies presents a previously unforeseen problem concerning the integration of form and motion. This issue is examined in greater detail in the following chapter.

The difference in the level of the direction sensitivity index between the conjoint form and motion selective cell population and the direction but not form selective cell population is predictable. The basis of the argument is found in the response characteristics of the form only selective cells and the directionally only selective cells in area STPa. The average tuning curve for direction falls slightly below the spontaneous firing rate (the mean direction index of cells selective for motion independent of form is 1.01, indicating that there was, on average, no response to motion in the null direction, Chapter 5). The average tuning curves of STPa cells selective for perspective view of the head and body remains above spontaneous activity throughout the possible 360 degrees of rotation (Perrett et al. 1991). This implies that cells receiving input from form selective cells within STPa will receive some form related input, regardless of the form of the stimulus. Therefore it is not surprising that a drop in the direction index is seen in cells selective for conjoint form and motion compared to cells selective for direction only since some "form" related input would be present from the non-preferred view.

Indeed, the estimated mean of I_d for the cells sensitive to direction independent of form (1.01) and the mean of I_v of cells sensitive to static bodies (0.66). If a non-linearity gain of 3.0 is applied only when both the preferred view and direction are present, the estimated mean I_d for cells selective for form and motion is 0.83, and the estimated mean I_v is 0.78 (see chapter Appendix). Note that this model predicts not only the drop in I_d but also the small increase in I_v. These are the values actually obtained, again suggesting that the input form and

motion signals to cells selectively responsive to the conjoint information are independent. Thus the behaviour of conjoint form and motion sensitive cells can be modelled as summing two independent inputs and applying an gain when the sum exceeds threshold (here threshold is greater than 1.34).

Appendix

Explaining the Discrimination Indices

For "View only" cells

$I_v = 0.66$

I_v normalized so that

response to preferred view = 1.0

therefore

response to opposite view = 0.34

For "Motion only" cells

$I_d = 1.01$

I_d normalized so that

response to preferred direction = 1.0

therefore

response to opposite direction = -0.01

The Model

"Form and motion" cells receive inputs that are the same as those into "Form only" and "Motion only" cells.

"Form and motion" cells have a non-linearity factor of 3.0 which acts only once activity levels exceed 1.5. At other times there is linear summation (i.e. The non-linearity is thresholded).

What does the model predict for I_d and I_v of "Form and motion cells ?

Sum of preferred view and direction

1.0 + 1.0 = 2.0

Non-linearity

2.0 * 3.0 = 6.0

Response to preferred view and direction = 6.0

Sum of preferred view and opposite direction

1.0 + -0.01 = 0.99

Sum of opposite view and preferred direction

0.34 + 1.0 = 1.34

Predicted I_d (Pref - Opp Dir)/Pref

(6.0 - 0.99)/6.0 = 0.835

(actual value = 0.83)

Predicted I_v (Pref - Opp View)/Pref

(6.0 - 1.34)/6.0 = 0.777

(actual value = 0.78)

CHAPTER 9

INTEGRATION OF FORM AND MOTION IN THE ANTERIOR SUPERIOR POLYSENSORY AREA (STPa) OF THE MACAQUE MONKEY III. TEMPORAL ASPECTS OF FORM AND MOTION INFORMATION INTEGRATION

(Oram & Perrett, 1995c, *J. Neurophysiol.* in submission)

INTRODUCTION

Visual processing in the primate brain is thought to occur along two largely separate pathways, the ventral, "what" or "form" pathway and the dorsal, "motion", "where" or "how" pathway (Newcombe and Russell 1969; Ungerleider and Mishkin 1982; De Yoe and Van Essen 1988; Felleman and Van Essen 1991; Young 1992; Goodale and Milner 1992; Milner and Goodale 1993). The anterior superior temporal polysensory area (STPa) of the macaque temporal lobe is one of the few visual cortical areas where the dorsal (motion) and ventral (form) pathways converge (Felleman and Van Essen 1991; Young 1992). Given the convergence of the two visual pathways in area STPa, it is not surprising that this cortical area is known to contain cells which are responsive to conjoint visual form and motion information (Bruce et al. 1981; Perrett et al. 1985a,b, 1989, 1990; see chapters 7 and 8).

One sub-population of these STPa cells responds selectively to the sight of bodies walking (or translating) in one direction and seen in one view (e.g. left profile go left). Evidence suggests that such STPa cells receive form and motion information from independent sources (see chapter 8). Further, the sources of the

form and motion information appears to be bound such that the motion signal must be derived from the moving body and not simply from the juxtaposition of these two visual cues (e.g. static body with the background moving in the preferred direction, chapter 7). A simple non-linear model suggested that the inputs to cells conjointly sensitive to form and motion were similar to the form input to cells selective for body view independent of motion and the motion input into cells selective for direction of motion independent of form (chapter 8).

Analysis of latency of three STPa cell populations ("form only", "motion only" and "form and motion") indicated that directionally sensitive cells and conjoint form and motion sensitive cells responded at a shorter latency than static form sensitive cells (chapter 8). Previous studies suggested that the form input to area STPa was as fast as possible (chapter 3; Oram and Perrett 1992; see also Thorpe and Imbert 1989), and also that the motion input to area STPa also arrived as fast as possible (chapter 5; Oram et al. 1993). If this is the case, then the differential latencies of the three populations of cells suggests that, for the early part of the response, conjoint form and motion sensitive cells would contain no form information. This non-discriminatory period would be expected particularly for conjoint form and motion sensitive cells with early latency response onsets.

The potential asynchrony of information arrival times (chapter 8) has implications for the integration of these two types of visual information. One current theory as to how the brain integrates information involves synchrony of activity in cells coding the same visual item (e.g. Singer 1993). It has previously been argued that synchronous arrival of multiple inputs appears to be an important feature of information processing in the ventral "form processing" pathway (Gochin et al. 1991; Geisler et al. 1991; Oram & Perrett 1992, 1994a; chapters 1,3). However, the role of synchrony for integrating form and motion remains unclear, and indeed seems to be challenged by the differential response latencies of "form only", "motion only" and "form and motion" cell populations in area STPa.

The observation of differential response onsets does not exclude the possibility of synchronization of inputs if there exists a oscillatory "carrier wave" which can time- or phase-lock inputs. Inputs from one source arriving synchronously or in phase with an oscillatory wave of excitation from a second source could be bound with each other. Such oscillations are known to exist in striate cortex of the cat (Gray et al. 1989; Engel et al. 1992a,b; Singer 1990a,b) and have also been found in the motion processing area MT of the macaque (Kreiter and Singer 1992). Oscillations do not appear to be present in inferotemporal (IT) cortex (Young et al. 1992), and their possible functional role is still unclear (Young et al. 1992; Tovee and Rolls 1992; Ghose and Freeman 1992; Singer 1993). The data to date would suggest that the ventral pathway (including the IT areas and area STPa) may use synchrony of inputs (see Gochin et al. 1991; Oram & Perrett 1992; chapter 3) for binding without oscillating carrier waves. The dorsal pathway (including area MT), may use oscillatory activity (Kreiter and Singer 1992) to bind motion information about contours of the same object.

The notion of a carrier wave to enable synchronization of inputs with differential arrival times is further complicated by the observation that information is reflected in the temporal modulation over hundreds of milliseconds of cell firing rates (Richmond et al. 1987; Richmond and Optican 1987; Optican and Richmond 1987; Gawne et al. 1991a,b; McClurkin et al. 1991a,b; Eskandar et al. 1992a,b). It is not clear whether temporal modulation of firing rate, if present (see Tovee et al. 1993), reflects information carried by the neuronal signals (feedforward) or whether such modulation is a result of feedback or lateral interactions between cells (chapters 1,3; Oram & Perrett 1994a). Such modulation could of course occur without disrupting any "carrier wave". Indeed, the frequency of the oscillations appears to be in the 40-90 Hz range in prestriate cortex, which is some two to five times higher than the observed modulations in IT cell responses.

In this article the temporal aspects of information arrival in the responses of cells sensitive to conjoint form and motion are considered. While there is insufficient data from single cells to directly investigate possible oscillatory activity, the aim of the present study was to investigate the extent of the possible asynchrony in the emergence of information about form and direction within STPa cell responses.

METHODS

Details of the recording methods and stimuli are given in the previous papers (chapters 7, 8). Briefly, standard techniques were used to record the activity of single cells in area STPa of the awake behaving macaque. While the monkey performed a colour discrimination task, images of walking bodies and control objects moving in up to four directions were presented. The bodies were either walking compatibly (following one's nose) or incompatibly (walking backwards) to the left, right towards and away from the subject. The neural spikes were converted to TTL signals and these were used to form peri-stimulus time histograms (PSTHs) with 250 bins. The PSTHs had a 1 second post-stimulus period and either 200 or 250 ms pre-stimulus sample periods. Eye movements were recorded using the sample period and stored with the spike data for each trial.

Data analysis

The dynamics of neural responses (the subject of the present study) were analysed only for cells that were classified as viewer-centred and unimodal (see Chapter 7). For each cell, responses to body-view and direction combinations were categorized into three groups: Preferred (response to the most effective body-view and direction combination), the Opposite View (Opp View, response

to the opposite body-view moving in the preferred direction) and the Opposite Direction (Opp Dir, response to the preferred body-view moving in the opposite direction). This grouping allowed the response levels to be compared across different cells when either the form (body view) of the stimulus or the direction of motion was changed independently of the particular direction of motion and view preferred by the individual cells. The response magnitude and temporal measures described in chapter 3 were calculated from each cell PSTH.

The discrimination between the Preferred and Opposite View categories was calculated as $100 * (R_P - R_{Ov}) / R_P$, where R_P is the mean response level above S/A in the Preferred response category and R_{Ov} is the mean response level above S/A in the Opposite View response category. If the Preferred response was within 1 spike per second of S/A the discrimination measure was not calculated. Similar discrimination measures were calculated using the Opposite Direction response category ($(R_P - R_{Od}) / R_P$, where R_{Od} is the mean response level above S/A in the Opposite Direction category).

Population PSTH response profiles in the three categories (Preferred, Opposite View and Opposite Direction) were generated (see Chapter 3 for method). To allow for combination of cell responses collected with slightly different time bins (4.8, 5.0 and 5.2 ms), a spike density function was obtained by convolution of the raw spike train data with a 2 ms Gaussian. This narrow Gaussian was used so as to minimalize temporal distortions due to spikes from one bin influencing data in other bins. Average responses (see Chapter 3) were also calculated for the Preferred, Opposite View and Opposite Direction response categories. Statistical efficiency of discrimination was determined using 2-way ANOVA (see Chapter 3) after Gaussian convolution of non-normalized response magnitudes of each cell.

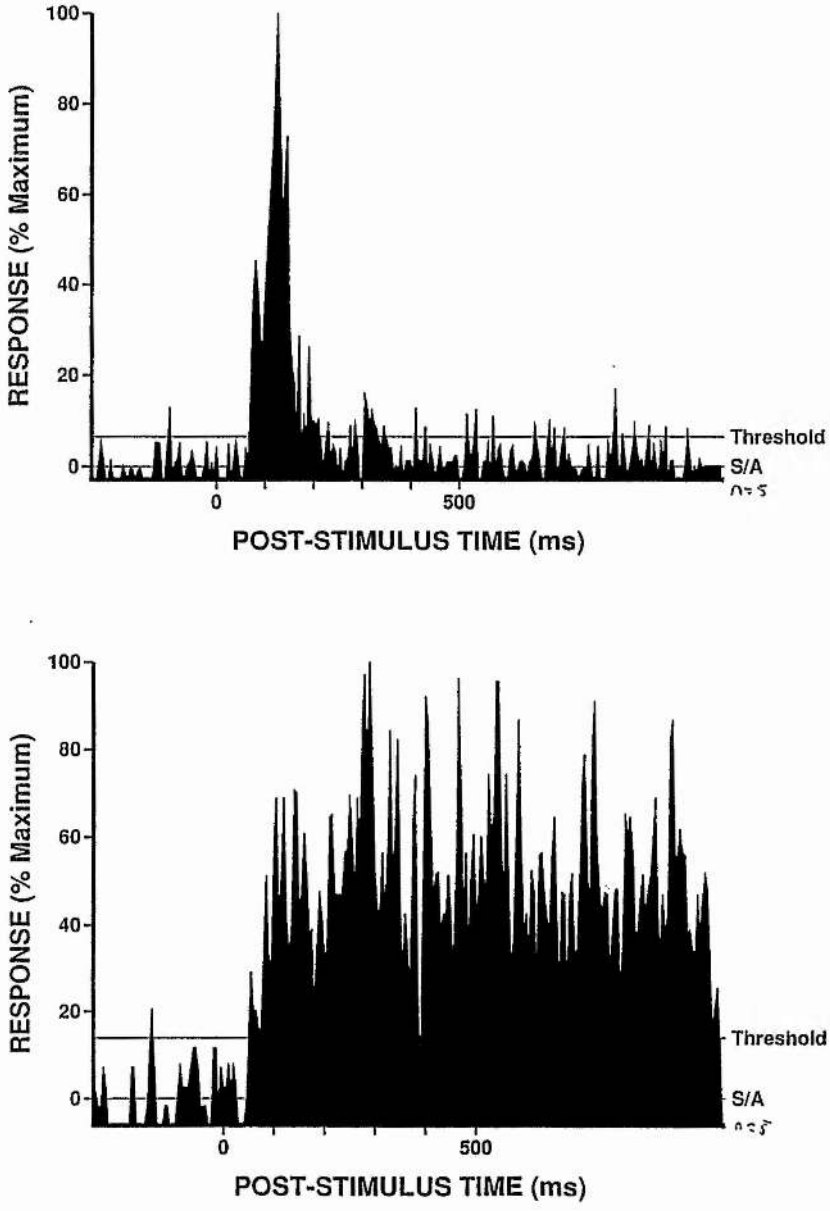


FIGURE 9.1 TRANSIENT AND SUSTAINED RESPONSES OF STPA NEURONS TO CONJOINT FORM AND MOTION STIMULI. Responses of single cells to effective form and motion stimuli. S/A = spontaneous activity, Threshold = 95% confidence limit of pre-stimulus activity. **UPPER: TRANSIENT RESPONSE.** The response to the effective stimuli produced a large transient burst of activity which lasted for some 150 ms before decaying. Note that the activity level at the end of the sample period is still slightly elevated compared to the pre-stimulus activity level. **LOWER: SUSTAINED RESPONSE.** A second cell produced a rapid increase in firing rate to effective stimuli. For this cell the response level was maintained throughout the sample period.

RESULTS

Neural recordings from three monkeys were used (*Macaca mulatta*, 2 male D, H wt. 5-8 kg, 1 female J wt. 4-7 kg from a UK registered breeding colony). 125 cells were classified as having selective uni-modal responses to the sight of moving bodies. This was approximately 2% of the total number of cells recorded in the STS upper bank (see chapter 7). Data were available in a suitable form for analysis of response time course for 43 of these cells.

To investigate the possibility of differences in the response measures between monkeys, 1-way ANOVAs were performed on all parameters listed in Table 9.1. The results indicated no significant difference between monkeys for all parameters ($P > 0.08$, all comparisons) except one out of 30 tested (2nd 100 ms Direction Discrimination, $F_{[2,40]} = 4.4$, $p < 0.02$). As the ordering of the mean Direction discrimination values showed no consistent pattern (i.e. rank ordering of Direction discrimination means for peak time = D,H,J, for 1st 100 ms = D,J,H, for 2nd 100 ms ordering = J,D,H) this was regarded as a Type 1 error. Parameter measurements for individual cells were therefore pooled across monkeys.

Time-course of response to effective stimuli

The measured response parameters for the Preferred category of response are listed in Table 9.1. The response of the cell population showed a rapid and clear increase in firing rate to the preferred view and direction combination, followed by a slower decline in firing rate. Evidence of the rise and fall in firing rate during the response was seen for all cells except one. The rapid rise in firing rate was a feature across most cells (rise time ranging from 5 to 239 ms, mean 76 ms). For some cells the decay of firing rate was very rapid and complete, the initial part of the response being a short, intense, transitory burst. Such transient activity is illustrated for one cell in the upper part of Figure 9.1. After the initial peak in firing rate, the response rate of this cell remained slightly increased over

Table 9.1. Parameter Summaries

Parameter	Mean	Range	N
A. Summary of Preferred View and Direction category responses			
Timing (ms)			
Latency	108.68	45.0 - 183.6	44
Rise Time	75.57	5.0 - 239.2	44
1/2 Fall Time	36.76	20.0 - 98.8	44
Decay Time	69.53	15.0 - 535.0	41[a]
Duration	106.96	15.0 - 580.0	41[a]
Firing rates (Spikes/S)			
S/A	10.26	0.0 - 39.1	44
Peak	110.68	40.0 - 217.9	44
1st 100ms	55.27	4.0 - 139.7	44
2nd 100ms	48.43	16.0 - 137.7	44
5th 100ms	33.46	4.0 - 88.6	44
End 100ms	31.63	0.0 - 86.5	44
Normalized to Peak response magnitudes (%)[b]			
1st 100ms	44.16	7.8 - 72.0	44
2nd 100ms	38.65	4.9 - 76.0	44
5th 100ms	23.25	-3.3 - 67.1	44
End 100ms	22.49	-6.3 - 67.1	44
B. Summary of direction (Dir) and view discrimination measures (%)			
Peak Dir Disc	34.24	-96.1 - 107.8	43[c,d]
1st 100ms Dir Disc	53.06	-182.6 - 200.0	43[c,d]
2nd 100ms Dir Disc	73.71	-136.4 - 202.9	43[c,d]
5th 100ms Dir Disc	83.57	-80.6 - 187.5	42[c,d]
End 100ms Dir Disc	71.38	-52.6 - 185.0	40[c,d]
Peak View Disc	29.26	-23.3 - 78.7	44[d]
1st 100ms View Disc	52.99	-102.5 - 140.9	44[d]
2nd 100ms View Disc	71.73	-129.5 - 147.9	44[d]
5th 100ms View Disc	70.81	-442.9 - 740.0	43[c,d]
End 100ms View Disc	52.61	-135.5 - 317.5	41[c,d]

S/A, spontaneous activity. [a] For three cells decay time and duration measures were not detectable within the sample period. [b] Normalized response magnitudes were calculated as (firing rate - S/A)/(peak - S/A). [c] Discrimination measures were not calculated if the Preferred response level was within 1 spike/s of S/A. [d] Negative discrimination measures resulted when the firing rate in the non-preferred category was greater than the Preferred, or when the Preferred category response level was less than S/A but greater than the non-preferred category response level.

the S/A for the duration of the stimulus presentation. Most cells showed a slower rate of response decay but still showed evidence of an initial transient burst. For

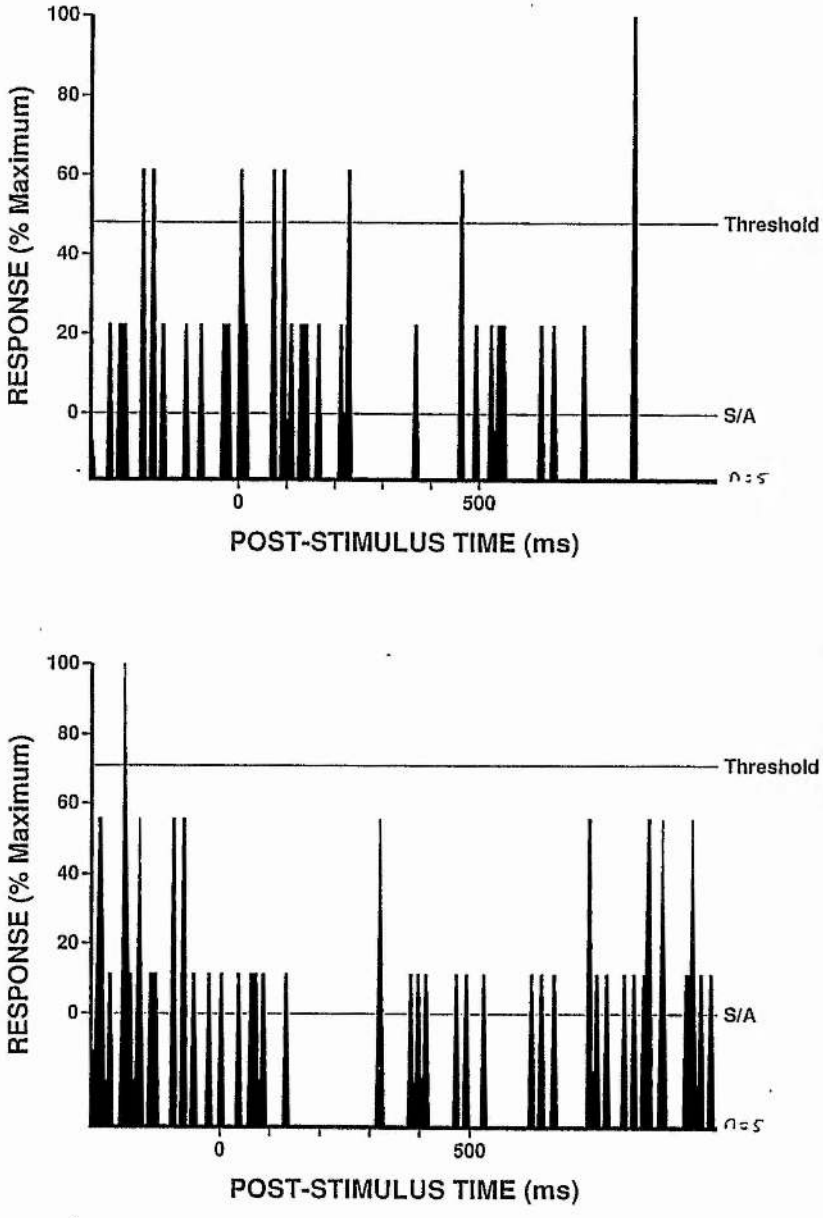


FIGURE 9.2 EVIDENCE OF INHIBITION TO INEFFECTIVE STIMULI. Example cell PSTH showing some evidence of inhibition to ineffective stimuli. S/A = spontaneous activity, Threshold = 95% confidence limit of pre-stimulus activity. **UPPER: INHIBITION OF OPPOSITE VIEW RESPONSE.** The response of one cell to the non-preferred view moving in the preferred direction is shown. There is evidence of inhibition from 230-475 ms post-stimulus onset. The response latency of this cell was 150 ms. **LOWER: INHIBITION OF OPPOSITE DIRECTION RESPONSE.** The response of a different cell to the preferred view moving in the non-preferred (opposite or null) direction is shown. There is evidence of inhibition from 100-300 ms post-stimulus onset. The latency of this cell was 85 ms.

one cell there was no evidence of a transitory burst, but rather the response was maintained at comparable levels throughout the sample period (Figure 9.1, lower). A gradation of decay time course between these two examples was found (decay time ranging from 15 to 535 ms). All cells showed evidence of a response maintained above spontaneous activity up to the end of the sampling period.

Preferred, Opposite View and Opposite Direction category responses

Comparison of the responses to three sets of stimuli allows the time course of information discrimination in cell responses under different conditions to be defined. As already noted, the Preferred responses show an initial burst and decay to a steady value which is greater than the pre-stimulus activity. In contrast there was a greater variety in shape of the Opposite view and Opposite Direction responses. At the population level there was a transient response similar in shape to the Preferred category but greatly reduced in magnitude. Only a few cells showed evidence of inhibition to the Opposite View or Opposite Direction stimuli in the first 50-200 ms of the response onset (see Figure 9.2). Inhibition, when seen, occurred some 20-50 ms after cell response latency. It is also relevant to note that the small amounts of inhibition were transitory, lasting 100-200 ms.

Although the population mean firing rate to the least effective view was never significantly below S/A, the negative firing rates indicate that inhibition was potentially present for some cells (e.g. Figure 9.2). Due to the low S/A of these cells (mean = 10 spikes/s), statistical detection of inhibition for any one cell was unlikely. Taking the response latency estimate from either the Preferred or All category estimate, the number of cells showing a firing rate to the Opposite View that was numerically below S/A was 0 for peak, 4 during the 1st 100 ms of response, 13 during the 2nd 100 ms of the response, 14 during the fifth and 10 during the final 100 ms period examined. The number of cells showing a firing rate to the Opposite Direction that was numerically below S/A was 1 for peak, 6 during the 1st 100 ms of response, 8 during the 2nd 100 ms of the response, 16

during the fifth and 12 during the final 100 ms period examined. All response levels at other times were above S/A. Cells responding at rates numerically less than S/A may not reflect real differences from S/A and may simply be due to sampling error. This was probable for several of the cells enumerated above, since differences between the estimated firing rate and S/A were often small (0.1 - 2.0 spikes/s). In summary there was little evidence of inhibition for non-preferred stimulus categories.

The time course and response amplitude parameters were compared across the different categories of response using a 2-way ANOVA analysis for the 42 cells where data was available for all three categories of response. The results of the comparisons between the three categories of responses are shown in Table 9.2. Not surprisingly, the average firing rate was significantly different across the three response categories during the early part of response. These response differences are maintained throughout the sample period (for at least 800 ms after response onset).

The latency estimates for the three categories varied significantly. This result, however, could be due to the weaker responses in the Opposite View and Opposite Direction categories. With a weaker response, it is more likely to over-estimate the response latency due to the characteristic of the square wave kernel used to create the bins. This is a likely explanation for this difference in response latencies between categories, given that the response onset to the non-preferred stimuli was co-incident with the response onset to the preferred stimulus for the calculated Average cell (see Figures 9.6,9.7,9.10).

The nature of the decay from the peak firing rate to the steady state firing rate also shows a significant difference between the response categories. Post-hoc testing indicates that the 1/2 fall time of the Preferred response category was significantly longer than that for the both the Opposite View and Opposite Direction categories of response ($P < 0.0005$ each comparison). A different rate

of response decay was also indicated by two other measures. The Decay Time (time

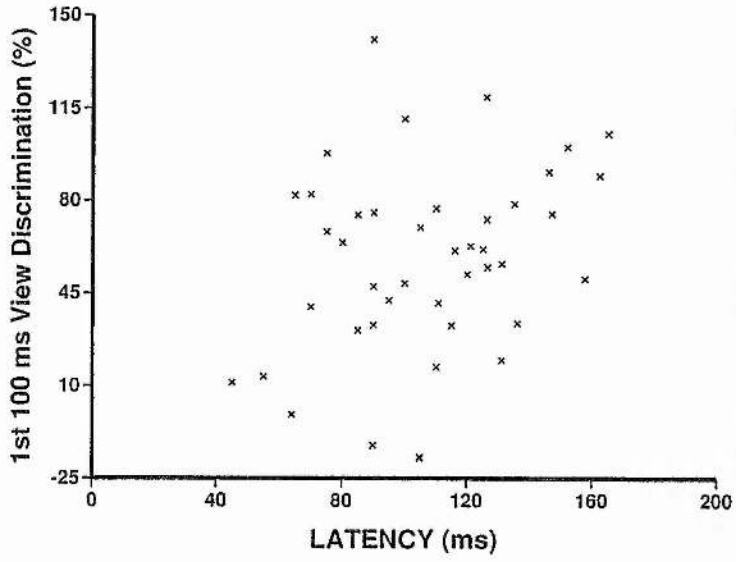
Table 9.2. Comparison of Preferred, Opposite Direction and Opposite View categories of response.

Parameter	PREF Mean	OPP DIR Mean	OPP VIEW Mean	F	df	p
Timing (ms)						
Latency	92.3	143.2	162.6	6.008	2/26	0.007
Rise Time	71.7	70.0	62.3	0.203	2/62	0.817
1/2 Fall Time	38.8	23.7	23.8	18.681	2/56	0.000
Decay Time	83.8	25.3	29.8	9.973	2/54	0.000
Duration	169.5	45.6	53.8	9.367	2/24	0.001
Firing rate (Spikes/S)						
S/A	10.3	9.6	9.9	0.555	2/84	0.576
Peak	111.1	73.9	80.2	39.737	2/84	0.000
1st 100ms	55.7	31.6	30.9	39.855	2/84	0.000
2nd 100ms	48.8	19.7	21.0	70.830	2/84	0.000
5th 100ms	33.5	14.8	18.0	37.722	2/84	0.000
End 100ms	31.6	15.9	20.6	22.730	2/84	0.000
Normalized to Peak Response magnitudes (%)						
1st 100ms	44.4	27.3	23.5	12.289	2/84	0.000
2nd 100ms	38.9	16.3	11.8	19.620	2/84	0.000
5th 100ms	23.2	6.1	7.7	14.493	2/84	0.000
End 100ms	22.3	9.4	10.5	11.299	2/84	0.000

Means for each parameter under each category are listed, with the resulting variance ratio (F), degrees of freedom (df) and probability (p) of the values being statistically indistinguishable.

from the peak of the response to a 3 time bin mean where the response is statistically equal to S/A) was also different across categories (Preferred > Opposite View, Opposite Direction, $P <= 0.001$ each comparison), as was the Response Duration ($P <= 0.001$ each comparison).

Further evidence for different rates of decay during the early part of the response (0-400 ms post-response onset) is apparent in Table 9.2 from the comparison of normalized response magnitude (expressed as a percentage of the peak response in each category). This latter calculation shows that the rate of



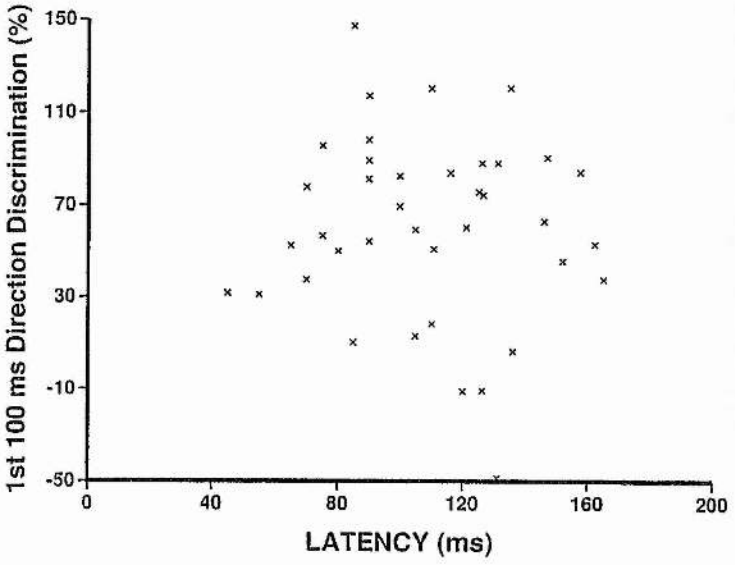


FIGURE 9.3 SCATTER PLOT OF LATENCY AGAINST DISCRIMINATION. The discrimination indices were assessed over the first 100 ms of the cell response. **UPPER: VIEW DISCRIMINATION.** Each cell (n=43) is represented by a single cross. Discrimination (I_V) is expressed as 100 * (Pref - Opp View) / (Pref - S/A). See text for details. There was a clear trend towards correlation between the two measures ($r_{[41]} = 0.291$, $p = 0.059$). **LOWER: DIRECTION DISCRIMINATION.** Each cell (n=42) is represented by a single cross. Discrimination (I_d) is expressed as 100 * (Pref - Opp Direction) / (Pref - S/A). See text for details. Here there was no significant correlation between the two measures ($r_{[40]} = 0.205$, $p > 0.1$).

decline is not linearly related to the peak response rate for that category (otherwise there would be no difference in this measure). What is more, the ratio of the normalised firing rates between categories is maintained at comparable levels throughout the sample period. For example, the ratio of the Preferred and Opposite View normalised firing rates, is 0.53 during the 1st 100 ms and 0.47 during the final 100 ms of the sample period. Given the substantial drop in firing rates over this time period, this again suggests that the mechanism leading to the differential decay rates seen in the response levels does not reflect absolute firing rates, but rather relative firing rates. Effects such as this are consistent with lateral inhibition or some other form of competition between cells.

Relation of response and latency

In deriving the time course of responses for the Average Cell the synchronization of the onset of response of the different cells to the same moment in time could potentially confound the relative times when direction or view discrimination occurred within a sub-population of cells. If it were the case that cells with longer latencies showed increased direction discrimination within the initial period of the response then synchronization could artificially decrease the time when directional information became evident within the response. This potential artifact, however, does not appear to be a problem with the sample of cells studied here.

Table 9.3 shows the correlation of the response parameters for the individual cells with latency. Of particular relevance is the observation that the direction discrimination index showed no tendency to correlate with response latency (Figure 9.3, lower). In contrast to this, there was a consistent and nearly significant weak positive correlation between response latency and the view discrimination indices (Figure 9.3, upper). This means that cells with longer response latencies tended to show clearer view discrimination, whereas directional discrimination appears to be independent of latency in this population of cells.

Table 9.3. Parameter correlations with latency.

Parameter	r	df	p
Timing			
Rise Time (ms)	-0.038	42	0.8060
1/2 Fall Time (ms)	0.005	42	0.9726
Decay Time (ms)	-0.228	39	0.1518
Duration (ms)	-0.244	39	0.1235
Firing rates			
S/A (Spikes/s)	-0.086	42	0.5807
Peak (Spikes/s)	-0.248	42	0.1050
1st 100ms (Spikes/s)	-0.141	42	0.3616
2nd 100ms (Spikes/s)	-0.101	42	0.5148
5th 100ms (Spikes/s)	0.037	42	0.8108
End 100ms (Spikes/s)	-0.082	42	0.5969
Normalized response magnitudes			
1st 100ms : Peak (%)	-0.012	42	0.9384
2nd 100ms : Peak (%)	0.029	42	0.8499
5th 100ms : Peak (%)	0.278	42	0.0676
End 100ms : Peak (%)	0.094	42	0.5430
Discrimination measures			
Peak Dir Disc (%)	0.199	40	0.2065
1st 100ms Dir Disc (%)	0.205	40	0.1933
2nd 100ms Dir Disc (%)	-0.130	40	0.4130
5th 100ms Dir Disc (%)	-0.008	37	0.9633
End 100ms Dir Disc (%)	0.049	35	0.7743
Peak View Disc (%)	0.368	42	0.0140
1st 100ms View Disc (%)	0.291	41	0.0586
2nd 100ms View Disc (%)	0.315	41	0.0394
5th 100ms View Disc (%)	0.419	36	0.0087
End 100ms View Disc (%)	0.274	38	0.0867

Correlation coefficient (r), degrees of freedom (df), and probability are listed for each parameter. S/A, spontaneous activity. Discrimination measures were calculated from the data after removing outliers.

Population response profiles of Early and Late latency neurons

The correlation between response latency and view discrimination was further investigated by splitting the cells conjointly sensitive to view and direction into two sub-populations. The Early cells were all cells whose response latency was less than the population mean response latency. The Late cells were those cells with a response latency greater than the mean.

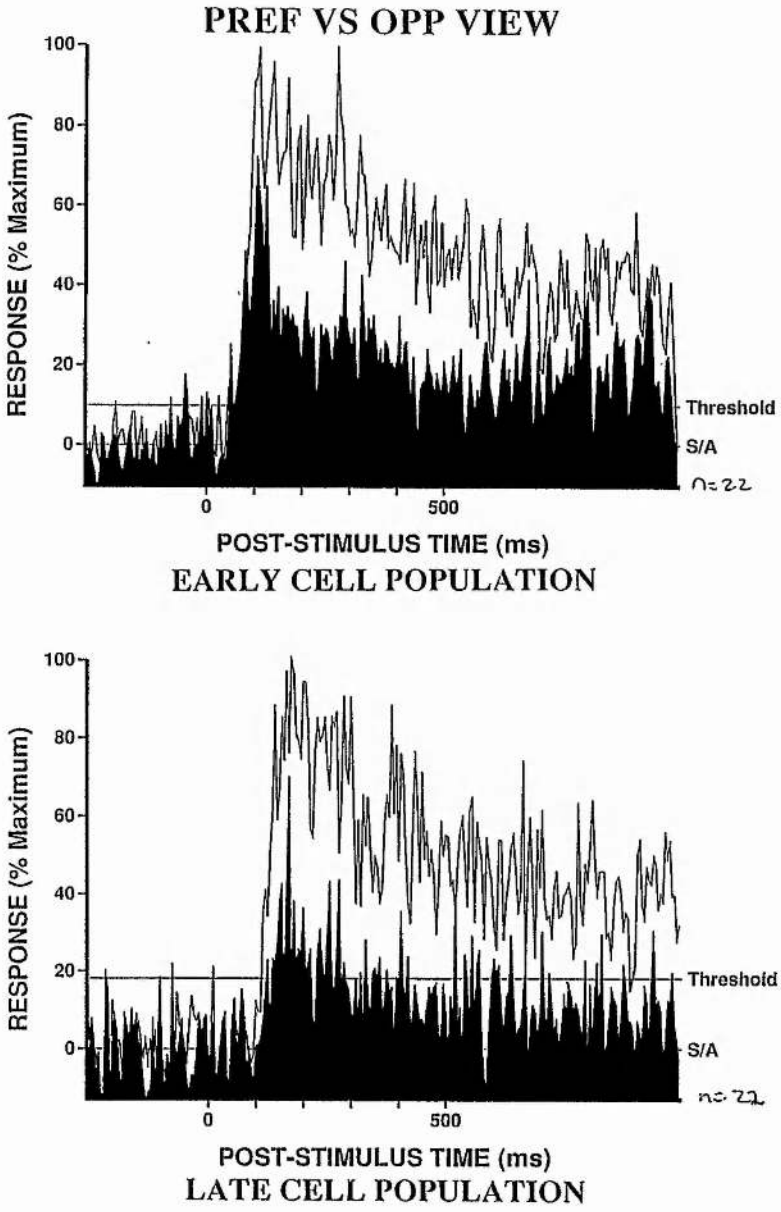


FIGURE 9.4 EARLY AND LATE CELL POPULATION RESPONSES TO THE PREFERRED AND OPPOSITE VIEW CATEGORIES. The normalized response profiles to the Preferred stimuli (clear) and the Opposite View stimuli (dark) are shown for the Early cell population (Upper) and Late cell population (Lower). Note the large response to the Opposite View stimuli is present for the Early but not Late cell populations.

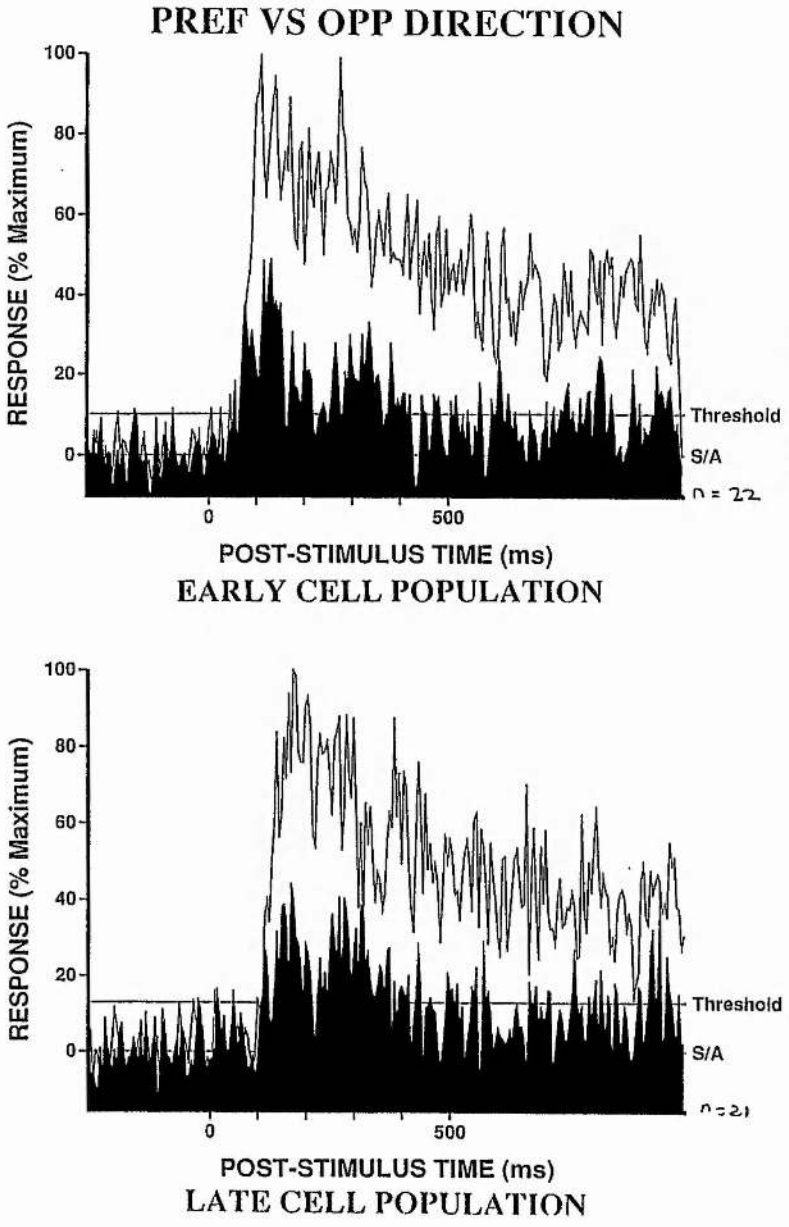


FIGURE 9.5 EARLY AND LATE CELL POPULATION RESPONSES TO THE PREFERRED AND OPPOSITE DIRECTION CATEGORIES. The normalized response profiles to the Preferred stimuli (clear) and the Opposite Direction stimuli (dark) are shown for the Early cell population (Upper) and Late cell population (Lower). Note the similarity of the two populations response profiles.

Early and Late cell population response profiles The effect of changing stimulus form (body view)

Figure 9.4 show the response profile of the Early and Late in the Preferred category and in the Opposite View category. As can be seen there is a substantial, non-discriminatory period at the beginning of the response of the Early cells (Figure 9.4, upper). This non-discriminatory period is not present between the response profiles of the Late cells (Figure 9.4, lower).

Early and Late cell population response profiles The effect of changing stimulus direction of motion

Figure 9.5 shows the response profiles comparing the Preferred and Opposite Direction category responses. It is clear that the differences in directional discrimination between the Early and Late cell populations is not as marked as that for view discrimination. This is consistent with the observation that the direction selective cell population within area STPa has an earlier average response latency than cells within the same cortical area sensitive to static body view (chapter 8).

Average Cell response profiles of Early and Late latency cells

In calculating the population level response profiles of the Early and Late cells, the process averaging across cells taking only a crude account of response latency. It is likely that, at the beginning of the response profile, only a few cells are active. Any assessment made using a population level analysis will therefore be subject to this potential error (see chapter 3; Oram & Perrett 1992). In order to control for this, the Average Early cell and the Average Late cell response profiles were calculated. In synchronizing the response latencies of all cells within each sub-population, the resultant Average Cell gives a more precise indication of response differences between the two sub-populations.

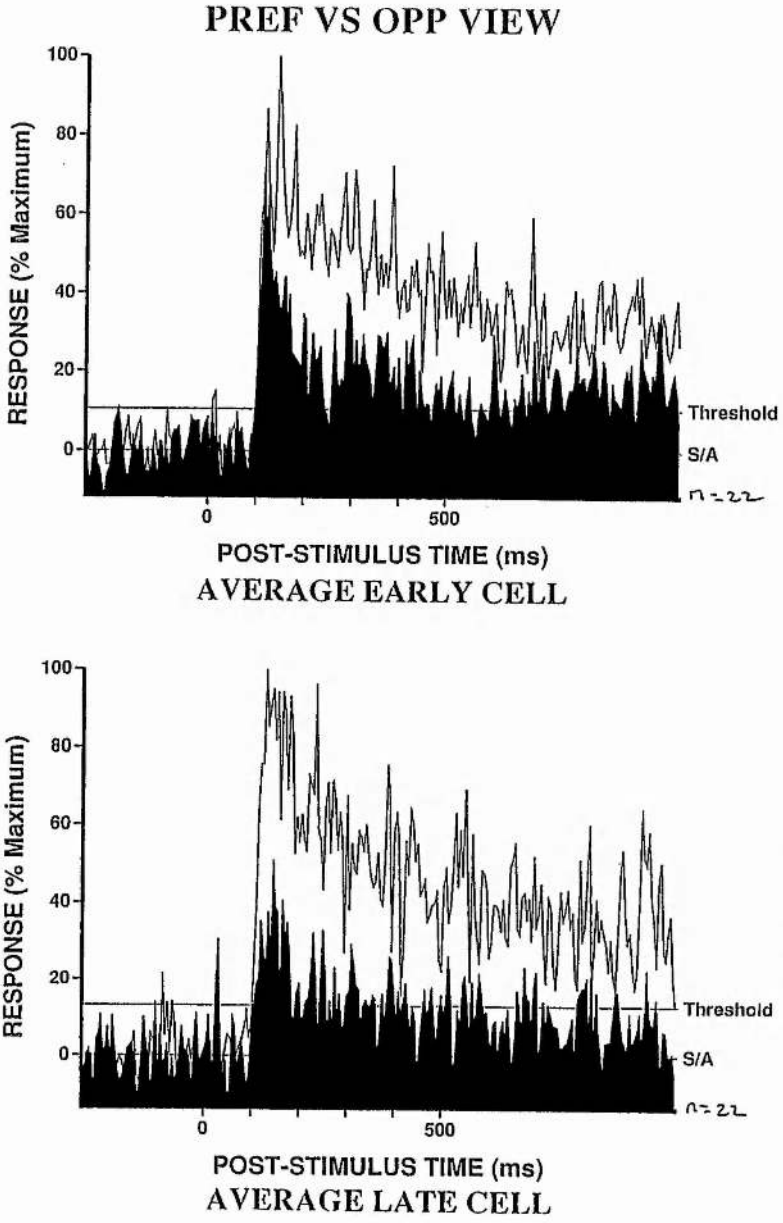


FIGURE 9.6 EARLY AND LATE AVERAGE CELL RESPONSES TO THE PREFERRED AND OPPOSITE VIEW CATEGORIES.

The normalized response profiles to the Preferred stimuli (clear) and the Opposite View stimuli (dark) are shown for the Average Early cell (Upper) and Average Late cell (Lower). The onsets of the cells were all aligned individually to 100 ms. The response onset at 100 ms for the Average cells indicates that the latency estimates were accurate. Note the large response to the Opposite View stimuli is present for the Average Early but not the Average Late cell.

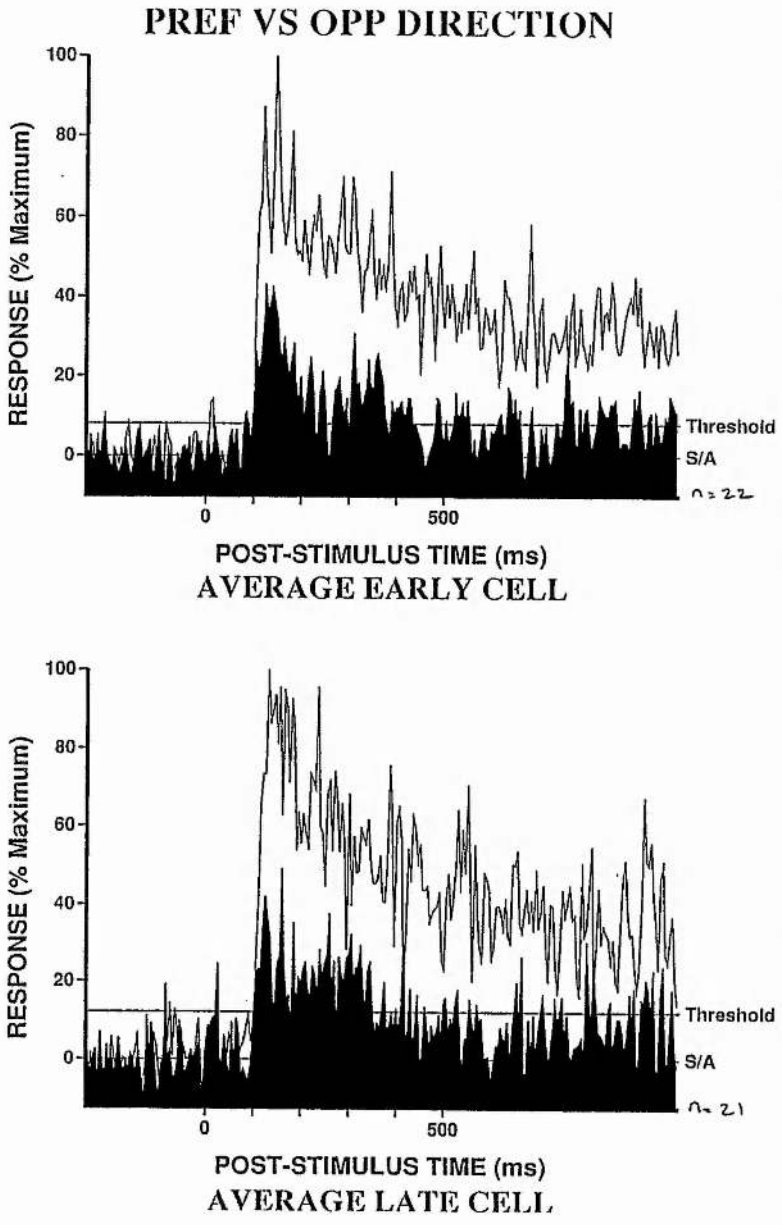


FIGURE 9.7. EARLY AND LATE AVERAGE CELL RESPONSES TO THE PREFERRED AND OPPOSITE DIRECTION CATEGORIES. The normalized response profiles to the Preferred stimuli (clear) and the Opposite Direction stimuli (dark) are shown for the Average Early cell (Upper) and Average Late cell (Lower). The onsets of the cells were all aligned individually to 100 ms. The response onset at 100 ms for the Average cells indicates that the latency estimates were accurate. Note the similarity between the response profiles of the Average Early and Average Late cells.

Average Early and Late Cell response profiles The effect of changing stimulus form

Figure 9.6 shows the response profiles of the Average Early and Late cells in the Preferred and Opposite View stimulus categories. The synchronization process shifted all cell latencies to 100 ms. The observation that this is indeed the "time" when the response profile reliably exceeds the threshold value indicates that the latency estimates for these cells was accurate. From the response profiles it is clear that the initial response period of the Early cells (Figure 9.6 upper) contains no view information. The Average Late cell, however, shows a large difference between the Preferred and Opposite View category response profiles immediately after response onset (within 10 ms).

Average Early and Late Cell response profiles The effect of changing stimulus direction

The Average Early and Late cell response profiles to the Preferred and Opposite Direction categories are shown in Figure 9.7. The difference between the Average Early and Late cell response profiles during the initial part of the response is small as both pairs of response profiles diverge with an equivalent time course.

Average Early and Late Cells The discrimination of stimulus form

A 2-way ANOVA (fixed factor = response category, random factor = cells) was performed between the Preferred and Opposite View categories of response for each of the 250 five ms time bins. The resulting F ratio values, where the rank ordering of the responses was Preferred > Opposite View, are plotted against time in Figure 9.8 upper for the Early latency sub-population. The lower part of the figure shows the analogous plot for comparison of the responses of the Late latency sub-population of cells. There is a period of approximately 50 ms during which the Average Early cell shows no view discrimination (Figure 9.8

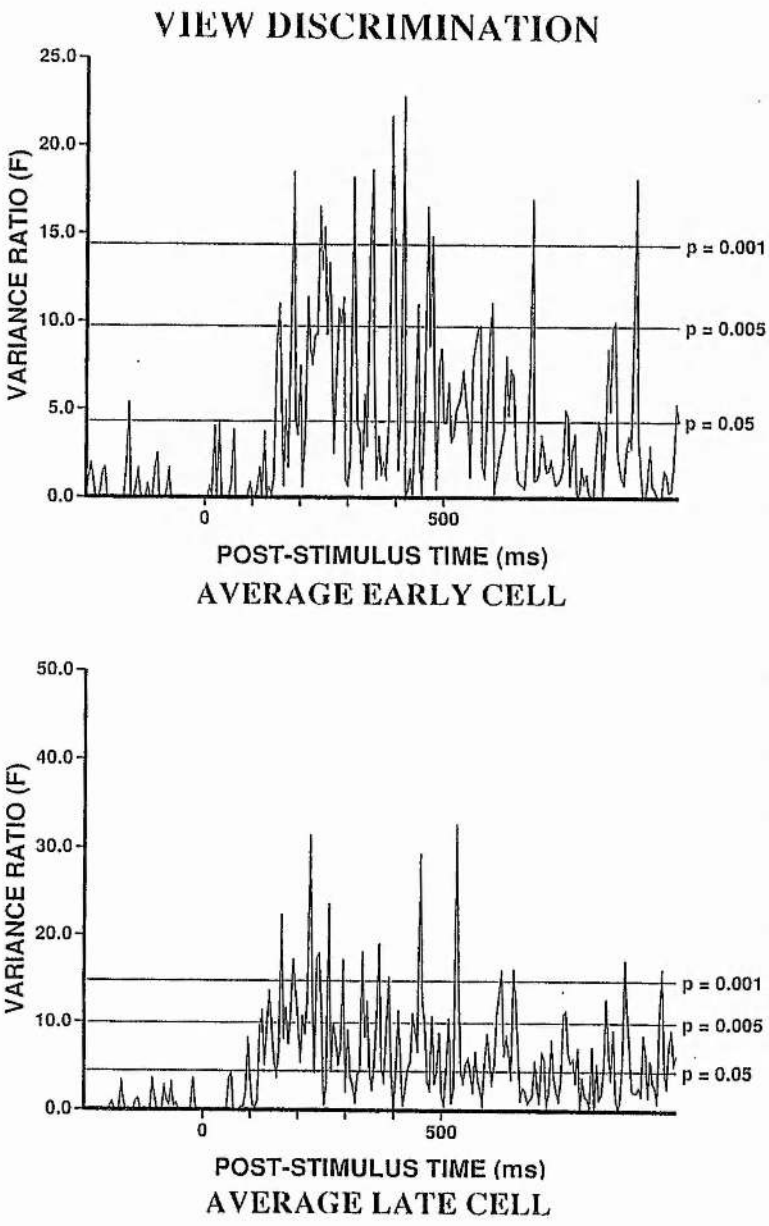


FIGURE 9.8 EARLY AND LATE AVERAGE CELL DISCRIMINATION BETWEEN THE PREFERRED AND OPPOSITE VIEW CATEGORIES. The variance ratio calculated between the cells' responses to the Preferred and Opposite View categories are plotted every 5 ms. **UPPER: AVERAGE EARLY CELL VIEW DISCRIMINATION.** Discrimination occurs at the 150 ms mark. The response onset (see Figure 9.7, upper) was set at 100 ms. **LOWER: AVERAGE LATE CELL VIEW DISCRIMINATION.** Discrimination occurs at the 120 ms mark. The response onset (see Figure 9.7, lower) was set at 100 ms. [The first peak at 95 ms occured before the response onset and is taken as a Type 1 error.]

upper). In contrast, the Average Late cell shows view discrimination within 20 ms of the response onset. To clarify this difference Table 9.4 gives the statistical assessment of view discrimination for successive time bins relative to response onset for the Average Early and Late cells.

Table 9.4. View discrimination of the Average Early and Average Late Cells.

	Early Cell		*Late Cell*	
Time (ms)	$F_{[1,22]}$	p	$F_{[1,20]}$	p
0	0.086	0.7717	3.276	0.0854
5	0.163	0.6904	0.478	0.4975
10	0.421	0.5232	0.339	0.5669
15	1.709	0.2046	0.989	0.3319
20	0.244	0.6261	9.492	0.0059 **
25	3.742	0.0660	11.171	0.0032 ***
30	0.059	0.8098	4.975	0.0373 *
35	0.546	0.4676	9.390	0.0061 **
40	0.200	0.6593	13.790	0.0014 ***
45	1.093	0.3072	9.750	0.0054 **
50	8.744	0.0073 **	5.484	0.0296 *
55	11.082	0.0030 ***	3.851	0.0638
60	5.918	0.0236 *	6.784	0.0170 *

The variance ratio (F) and the associated probability (p) are listed for view discrimination (Preferred vs Opposite View) for the Average Early and Late cells. The Average Early Cell was calculated from cells whose response latency was below the population mean, the Average Late cell was calculated from those cells whose response latency was greater than the population mean. Time gives the time after the response onset. Significance is indicated by the * (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$).

Discrimination analysis was also performed using transformed data to correct for the standard deviation being proportional to the mean (logarithmic transform). Variance in firing rate can be correlated with the response strength (Vogels and Orban 1991). A suitable transformation [(firing rate + 1)$^{1/2}$, Snedecor and Cochran 1980] was performed to allow for this relationship, or a Poisson distribution underlying firing rate, which might invalidate ANOVA assessment of discrimination. The results using these transforms showed the same pattern, with a 50 ms delay before view discrimination was apparent for the Average Early cell and a 20 ms lag for body view discrimination to become apparent between the responses of the Average Late cell.

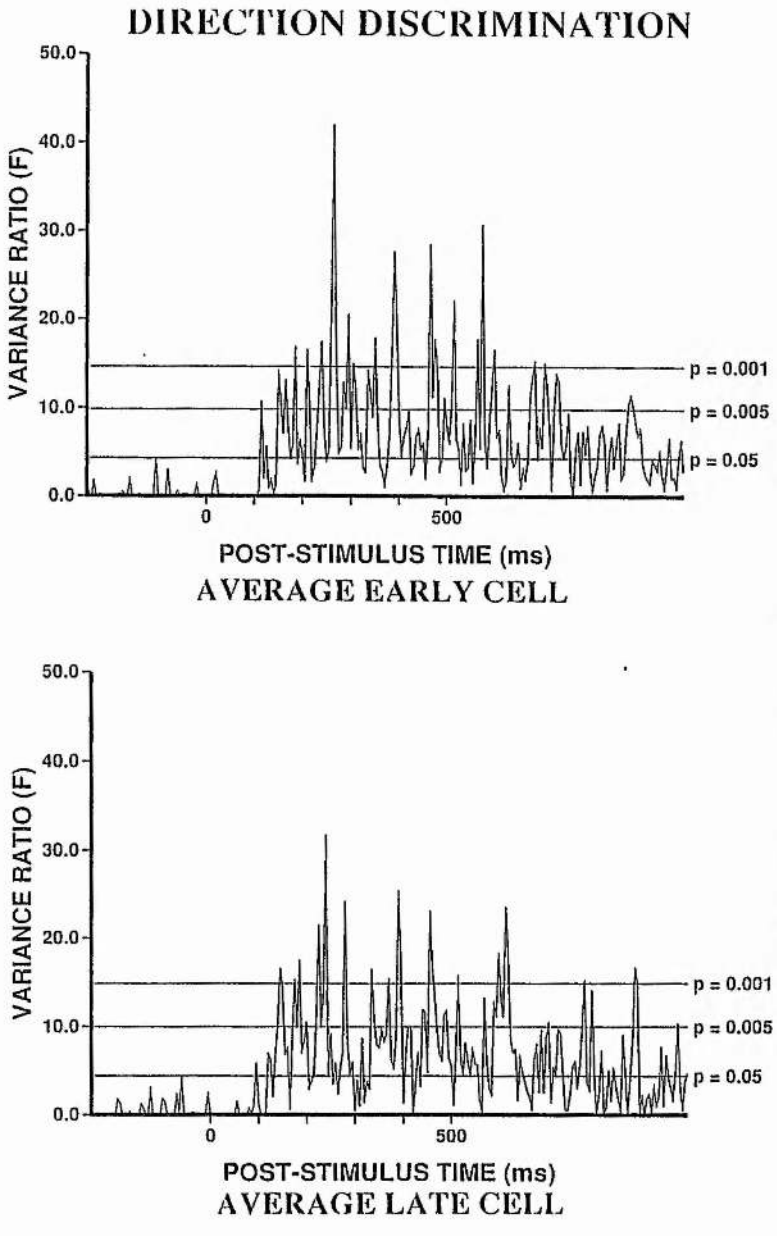


FIGURE 9.9. EARLY AND LATE AVERAGE CELL DISCRIMINATION BETWEEN THE PREFERRED AND OPPOSITE DIRECTION CATEGORIES. The variance ratio calculated between the cells' responses to the Preferred and Opposite Direction categories are plotted every 5 ms. **UPPER: AVERAGE EARLY CELL DIRECTION DISCRIMINATION.** Discrimination occurs at the 115 ms mark. The response onset (see Figure 9.8, upper) was set at 100 ms. **LOWER: AVERAGE LATE CELL DIRECTION DISCRIMINATION.** Discrimination occurs at the 120 ms mark. The response onset (see Figure 9.8, lower) was set at 100 ms. [The first peak at 95 ms occured before response onset and is taken as a Type 1 error.]

Table 9.5. Direction discrimination of the Average Early and Average Late Cells.

	Early Cell		*Late Cell*	
Time (ms)	$F_{[1,21]}$	p	$F_{[1,20]}$	p
0	0.041	0.8418	1.338	0.2611
5	0.565	0.4608	0.002	0.9682
10	0.736	0.4005	0.133	0.7190
15	10.816	0.0035 ***	0.292	0.5950
20	1.903	0.1823	6.854	0.0165 *
25	5.711	0.0263 *	6.192	0.0218 *
30	1.063	0.3142	2.005	0.1722
35	1.892	0.1835	6.982	0.0156 *
40	0.315	0.5804	10.137	0.0047 ***
45	1.210	0.2839	16.581	0.0006 ****
50	14.246	0.0011 ***	14.331	0.0012 ***
55	10.717	0.0036 ***	6.861	0.0164 *
60	7.053	0.0148 **	7.499	0.0127 *

The variance ratio (F) and the associated probability (p) are listed for direction discrimination (Preferred vs Opposite Direction). The Average Early Cell was calculated from cells whose response latency was below the population mean, the Average Late cell was calculated from those cells whose response latency was greater than the population mean. Time gives the time after the response onset. Significance is indicated by the * (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, **** $p < 0.001$).

Average Early and Late Cells The discrimination of stimulus direction

Figure 9.9 and Table 9.5 show the values of direction discrimination obtained from the Average Early and Late cells. From these there is very little difference between the Average Early and Late latency cells in the time from response onset to the onset of statistically significant direction discrimination (15 ms and 20 ms). The same lags were found using square-root or logarithmic transformed data.

Given the similarity of the Average Early and Late cells for direction discrimination, and the fact that there was no correlation between response latency and the direction discrimination indices, all cells were combined to produce a more accurate description of direction discrimination shown by the Average cell conjointly sensitive to form and direction. The response profile and discrimination of this Average Cell are shown in figure 9.10.

Even with the combined data from 43 cells, it is clear that for the Average Cell there is a short lag of 10 ms from response onset (aligned to the 100 ms

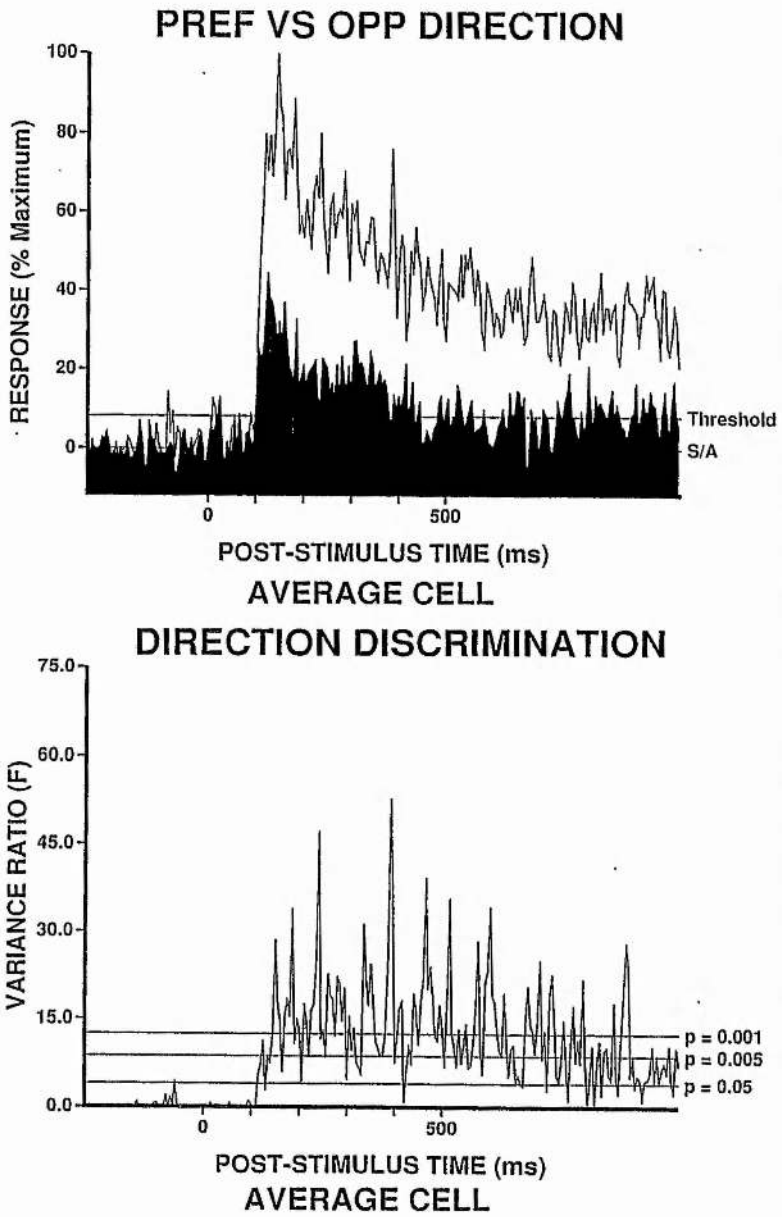


FIGURE 9.10. THE AVERAGE CELL: PREFERRED AND OPPOSITE DIRECTION RESPONSE PROFILES AND DISCRIMINATION. The data comparing the Early and Late cell Preferred and Opposite Direction category responses were collapsed. **UPPER: AVERAGE CELL RESPONSE PROFILES.** The normalized response to the Preferred stimuli (clear) and Opposite Direction stimuli (dark) are plotted. **LOWER: AVERAGE CELL DIRECTION DISCRIMINATION.** The variance ratio (F) between the Preferred and Opposite Direction stimuli (dark) are plotted every 5 ms. The response onset is aligned to 100 ms. Discrimination becomes evident at 110 ms.

mark) to the onset of significant direction discrimination. In other words the first 10 ms of the response in non-discriminatory for direction.

Response properties of STPa cells to effective stimuli

The responses of the cells selective for conjoint form and motion information showed typical response profiles to the preferred stimuli. The initial phase of the response was a fast increase in firing rate, followed by a slow decay to a steady firing rate. This basic pattern of fast rise and slower decay has been observed in "form only" cells (those cells selective for static form information, chapter 3; Oram and Perrett 1992) and "motion only" cells (cells selective for direction of motion independent of form, chapter 5; Oram et al. 1993). To allow comparison between these populations, the means of all response measures for these three populations of cells are given in table 9.6.

As is evident from table 9.6, the temporal aspects of the responses between the three cell populations are similar in many aspects except for the response latency ($p = 0.001$). As is found in striate cortex, the cells showing sensitivity to motion can be distinguished by the earlier response latency. There is also a tendency for cells sensitive to "motion only" to show a slower drop from peak firing rate to half peak ($p = 0.002$). Interestingly this is not accompanied by an increase in the total time for the response to decay down towards threshold ($p > 0.1$) or in the total response duration ($p > 0.2$). Given the difference in latency of the form and the motion inputs, the temporal characteristics of the conjoint sensitive cell responses are predictable. For conjoint form and motion sensitive cells, the slowly decaying motion input is followed by a faster decaying form signal. STPa neurons selective for conjoint form and motion show a decay rate similar to "form only" cells ($p > 0.5$). The later arriving form input gives rise to

Table 9.6. Parameter Means of Form and Motion, Motion Only and Form Only cell populations.

Parameter	CELL TYPE Form + Motion	Motion only	Static View
Timing (ms)			
Latency* (ms)	98.9	90.9	119.1
Rise Time (ms)	75.6	69.4	58.2
1/2 Fall Time (ms)	36.8	59.0	40.0
Decay Time (ms)	69.5	134.4	93.4
Duration (ms)	107.0	168.6	112.5
Firing rates			
S/A (spikes/s)	10.3	11.4	8.6
Peak (spikes/s)	110.7	108.3	115.1
1st 100ms (spikes/s)	55.3	67.3	66.9
2nd 100ms (spikes/s)	48.4	53.1	48.1
5th 100ms (spikes/s)	33.5	31.9	28.5
End 100ms (spikes/s)	31.6	30.6	24.7

The parameter means for the timing and response magnitudes are shown for three STPa cell classes. Form + Motion = cell population sensitive to conjoint form (body view) and motion information. Motion only = cell population selective for direction of motion, independent of form (from chapter 5; Oram et al. 1993). Static View = cell population selective for static body view. * Latency estimates are taken from the earliest estimate from either the All or Preferred (most effective stimulus) response category.

the peak of the response and is the main contributor to the 1/2 fall time. In area STPa, it would appear that cells with form selectivity show a faster decay rate (1/2 fall time) than cells showing motion but not form selectivity. This is in contrast to the cells of striate cortex which can be divided into transient (motion selective) and sustained (orientation selective) responses.

The response magnitudes during all periods of the responses of cells selective for conjoint form and motion were similar ($p > 0.1$ at each time period examined) to those seen for static body view selective cells (chapter 3; Oram & Perrett 1992) and to cells sensitive to the direction of motion independent of form (chapter 5; Oram et al. 1993).

DISCUSSION

Differences between response categories

As seen with cells in area STPa sensitive to static form there was a faster decay in firing rate observed to less effective stimuli than to effective stimuli. This was observed in three timing measures: the 1/2 fall time, the decay time and the response duration. In all cases, the measure of the speed of response decay were statistically equivalent when either the body view (form) or direction of motion was changed from the optimal by 180 degrees. Further, changes of either body view or direction of motion produced equivalent changes in the response decay. These between stimulus effectiveness category differences in decay rate were reflected in the normalised response magnitudes. This indicates that the differences were not due to simply the differential firing rates.

The between-category differences in response decay rate could reflect interactions between STPa cells, feedback loops (involving either higher or earlier processing stages) or interactions between cells at earlier stages of processing. The present data do not allow distinction between these possible sources of the differential decay. The data are, however, consistent with the notion of extraction of the unaccounted information (Pece 1993; chapter 1; Oram & Perrett 1994a). Under this scheme, which has proved a useful technique in computational approaches to vision (Seibert and Waxman 1993), information regarding the differences between the input signal and stored exemplars is used to enable more precise categorisation of the input signal (chapter 1; Oram & Perrett 1994a).

The Early and Late Average cells

Examination of the response latencies of STPa cell populations indicates that inputs containing motion information arrive before inputs containing form information (chapter 5; Oram et al. 1993; chapter 8). Consistent with this is the observation that there was little difference between the Early and Late latency

cells when comparing the Average cell response profiles to the Preferred and Opposite Direction categories (Figure 9.7). This similarity was confirmed from the discrimination measure between these two response categories (Figure 9.9).

Studies of MT cells using apparent motion stimuli indicate that there is a non-discriminatory initial response to the first of the sequence of light points (Mikami et al. 1986a,b; Newsome et al. 1986). This suggests that the initial response period of direction selective cells in MT is non-discriminatory and non-directional. The same appears to be true in STPa cells sensitive for conjoint form and motion information, where the first 10 ms of the response is non-discriminatory (Figure 9.10). This is in marked contrast to cells in area STPa selective for static form, which show discrimination within the first 5 ms of response onset (chapter 3; Oram and Perrett 1992). The non-discriminatory period seen in the present study for direction in the conjoint form and direction selective cells is unlikely to be an artifact due to the number of cells used in the analysis: the investigation into form discrimination by Oram and Perrett (1992; see also chapter 3) used only 22 cells to calculate the Average cell response profile and discrimination, yet calculations for the Average cell shown in Figure 9.10 used 43 cells. Note also the increased significance level and reliability of the direction discrimination shown in Figure 9.10 compared to the direction discrimination shown for the Average Early or Late cell (Figure 9.9).

Given the difference in response latencies of cells selective to static form and cells selective for direction of motion, it would be expected that the initial response of cells with early response latencies would be driven by inputs from motion selective cells. The Average Early cell shows little difference during the early part of the response between the Preferred and Opposite View categories (Figure 9.6, upper). This is in contrast to the response magnitudes to the same categories for the Average Late cell (Figure 9.6, lower). The view discrimination profiles of the Average Early and Late cells (Figure 9.8) show a clear difference,

with the Average Early cell showing a 50 ms non-discriminatory period during the initial phase of the response.

It would appear that the motion inputs are directionally non-discriminatory for a period of approximately 10 ms. Inputs containing form information arrive in area STPa somewhat after the motion direction inputs. Estimates of response latencies of cells in STPa suggest that the delay is of the order of 20 ms (Table 9.6). Previous studies of STPa cells sensitive to form suggest that discrimination in these cells is present from response onset (chapters 1,3; Oram & Perrett 1992, 1994a). While form discrimination was not present at response onset in cells conjointly selective for form and motion, this does not mean that the form information, when it arrives, is non-discriminatory. The division into Early and Late responding cells was done by dividing the population studied about a latency of 100 ms. Given that the mean latency of cells selective for static form is 119 ms, it is likely that some "late" responding form and direction selective cells still might not have received their form inputs at response onset.

In summary area STPa receives a motion signal (presumably from the dorsal pathway) which is initially non-discriminatory for direction. About 20 ms later STPa cells receive an input (presumably from the ventral pathway) which discriminates form.

Role of attention in integrating form and motion information

Psychological models of integration of different stimulus attributes propose an underling attention system (e.g. Treisman & Gelade 1980). It is therefore relevant to consider the possible role of attention in the cell responses described here. If attention underlies the integration of form and motion the response profiles of the Early and Late cell populations would be equivalent. This is clearly not the case (Figures 9.4, 9.6). It is of course possible that attention effects do not play a part in the initial period of the response of the Early cells, and that the information is not bound until the attention system comes into play.

There are however reasons outlined below that make this role of attention unlikely.

Evidence from studies of inferotemporal neurons suggest that the effect of "concious" attention has a modulatory role rather than gating the responses. For example, the study of Chelazzi et al (1993) indicated that attention to one object type modulated receptive field size (see Desimone et al. 1990 for review) and while having a pronounced effect on response magnitude, this effect was not seen for some 100-200 ms after response onset. The use of a warning tone, as in the experiments report here, just prior to stimulus presentation can suppress cell pre-stimulus activity of cells in IT cortex (Sato et al. 1980). Relatively minor effects on STP and IT cell response magnitudes have been observed using a second concurrent visual stimulus or fixation spot; these include both decreases (Richmond et al. 1983) and increases in response rate (Moran and Desimone 1985; Richmond and Sato 1987; Sato 1988, 1989; Fuster 1990). Given that attention in the behavioural task was to an LED, attention effects may have caused a slight decrease in response magnitude to the test stimuli but other effects are unlikely.

While the LED colour discrimination task is relatively easy (lick reaction times 300-400 ms, Perrett and Oram, unpublished observations), the monkey maintained fixation on the LED during the first 200-400 ms of stimulus presentation (chapter 7). The period used to assess whether the cells of the present study showed conjoint selectivity was 100-350 ms. During this period the monkey was fixating the LED. As the monkeys were free to move their eyes at any time during the trial, this strongly suggests that the monkey was indeed attending the LED.

In summary, whatever the mechanism used to achieve conjoint form and motion integration, it is unlikely that there was "attention" monitoring of "what" and "where" signals. The responses observed were fast and indicated integration during the period when attention was directed at the LED, not at the stimuli. This

suggests that the mechanism used for integration of form and motion is an autonomic one based on cell and pathway properties themselves rather than an attention system.

Temporal aspects of information processing

It has previously been argued that the flow of visual information through the macaque brain is as fast as possible given the anatomical constraints, both for the ventral "form" pathway (Oram and Perrett 1992) and through the dorsal "motion" pathway (Oram et al. 1993). The data from the present study supports this idea. The response latency of cells selective for both form and motion is comparable to cells selective for direction but not form. Cells with early response latencies do not contain any form information during the initial 50 ms or so of the response. Form and direction selective cells responding with a longer latency show faster emergence of form information in their responses. As argued above, this is consistent with two types of visual information deriving separately from the two main cortical visual processing pathways.

Differential arrival times of spikes has been proposed an efficient method of coding information (Thorpe 1990). The present data indicate that responses to less effective stimuli have the same latency as responses to effective stimuli. The simultaneous response onset to effective and ineffective stimuli has also been found for form selective neurons in area STPa (Oram and Perrett 1992). It would therefore appear as if coding of visual pattern attributes in temporal cortex does not utilize differential spike arrival times.

It has been suggested that response oscillations and/or response synchronization could underlie the integration of different visual attributes (von der Malsberg 1988; von der Malsberg and Schneider 1986; Gray et al. 1989; Kreiter and Singer 1992; Engel et al. 1992a,b; Singer 1990, 1993; Singer et al. 1990). The present data challenge this idea. The response and discrimination profiles of cells responding to conjoint form and motion stimuli with early

response latencies (Figures 9.6-9.10) indicate differential arrival times of form and motion information. This suggests that synchrony does *not* play a role in the integration of form and motion in area STPa. Synchrony could of course be used to aid transmission within each pathway. The role of this synchronous activity, other than to allow the rapid transmission seen in both pathways (Oram and Perrett 1992; Oram et al. 1993; Oram and Perrett 1994a), is however unclear.

As synchrony does not appear to play a role in binding the possible role of temporal coherence (oscillations) is considered below. While there are connections between V4 and MT/FST, temporal coherence would have to be maintained through PIT, CIT and AIT of the ventral pathway and MST and STPp of the dorsal pathway. Given the differential transmission rates of these pathways and the difference in number of areas to reach STPa via the ventral and dorsal routes, it seems unlikely that temporal coherence of any waveform would be maintained accurately across the two systems.

It is important to note that the findings reported here do not mean that synchrony or oscillations are not used in the primate brain to bind information. The data presented here applies only to binding of form and motion: perceptually we bind many other features, such as position, colour, contours of the same object and so on. The results do, however, suggest that such a mechanism, if present, is not employed for the integration of form and motion.

Oscillations may be used to bind information between processing elements within the dorsal pathway. However, there is as yet no clear evidence that oscillations are present in the ventral pathway. The absence of oscillations in the ventral pathway would preclude their use in binding both within that pathway and between the two pathways.

The constant lag between arrival of form and motion information at any one cell selective for both attributes is also inconsistent with the use of temporal coherence between the two pathways as a basis for information binding. It might appear that oscillations can be used for binding despite a fixed delay across two

pathways, providing that the period of the oscillation is a multiple of the delay. This places a severe restriction on the use of different carrier frequencies between 40 and 100Hz to match a lag of 20ms. Effectively such a restriction means that different items could not be bound by different carrier frequencies. The availability of only one carrier frequency (50Hz) means that information form all objects in the scene would be bound together creating a multitude of illusory conjunctions.

CHAPTER 10

SUMMARY

In this chapter the major findings and conclusions are brought together and summarised. The first section in this chapter considers form processing (from chapters 1,3 and 4). Motion processing within area STPa is then summarized (chapters 5, 6). The third section looks at the integration of form and motion. Finally, further applications of "time course" analysis of neurobiological data are considered.

1. FORM PROCESSING

Speculative outline of visual form processing

Neurobiological data from the cerebral cortex of the macaque monkey suggest a model of object recognition which involves a series of four computational stages. These are executed in seven major hierarchically arranged areas of processing, each area with an input and an output layer of cells.

The first computational stage occurs within early visual cortex and involves the first two cortical areas (V1,V2). Here it appears that boundaries between image regions and logical groupings of local oriented image elements that 'belong' together are computed. These processes segregate image attributes which can then be treated as arising from the same object.

The next three visual cortical areas (V4, PIT, CIT) execute the second computational stage and display sensitivity to an ever increasing complexity and variety of visual shape features (e.g. T junctions, concentric rings, spotted triangle shape). The third stage of processing seems to utilize combinations of these shape

features to establish selectivity to object-feature instances (i.e. the approximate appearance of a small number of object attributes seen under particular viewing conditions). Cells in the area tolerate change in position but show only limited generalization for change in retinal size, orientation or perspective view.

The fourth proposed computational process occurs within the final cortical areas (including AIT and STPa) and gives rise to cell selectivity showing object constancy across size and orientation. This process probably occurs through pooling of the outputs of cells responsive to different instances of the same object view. Importantly, constancy across perspective view (i.e. the transition between viewer-centred and object-centred representation) does not seem to be completed except by a small percentage of cells in these areas.

It is argued that top-down influences, though poorly understood, may play a role in nulling image aspects that are predictable in appearance and/or not the object of attention such that only features containing relevant discriminatory information are further processed. This postulated role for feed-back connections would allow behavioural influence on the learning process. Such influence would allow the organism to adapt to its particular environment.

Synaptic changes encompassing various associative (e.g. Hebbian) and non-associative (e.g. decorrelating) procedures may allow cells throughout the stages of processing to become tuned to frequently experienced image attributes, shapes and objects. Associative learning procedures operating over short time periods may underlie the progressive generalization over changing viewing conditions. It is suggested that constancy across position, orientation, size and, finally, perspective view and object parts is established slowly as each area pools the appropriate outputs of the less specific cells in the preceding area.

The efficiency of biological form processing

In chapter 3 the speed at which form information becomes evident in STPa cell responses was examined. The results of this study indicated severe constraints on the nature of information flow through the ventral "form" pathway.

Measurements of the magnitude and time-course of response were made from 44 cells responsive to static head views at different levels of stimulus effectiveness. In this way responses to complex stimulus patterns evoking good, poor and mid-range responses could be compared across one form sensitive cell population. Cells exhibiting both good and poor initial discrimination between head views were found at short and long latencies; there was no correlation of any of the temporal response parameters measured with cell response latency.

The time course of the population response to the most effective stimuli showed a rapid increase to a peak firing rate (onset to peak, rise time = 58 ms) that was on average 115 spikes/s above spontaneous activity (S/A), followed by slower decay (decay time = 93 ms) to a maintained discharge rate (15% of the peak rate above S/A). Discrimination between responses to different head views exhibited by the population showed a sharp rise and reached highly significant levels within 25 ms after the population's response onset.

On average, activity in a single neuron (the Average Cell) rises to 44% of its peak response rate within 5 ms of the response onset. The Average Cell also showed exceptionally fast discrimination between views, significant within 5 ms of response onset. It is argued that the fast rise in firing rate, followed by a decay to a lower rate and the very fast emergence of discrimination are features of pattern processing present in real neural systems that are lacking in many processing models based on artificial networks of neuron-like elements, particularly those where discrimination relies on top-down and/or lateral competitive inhibition. It is concluded that the only way to account for the rapid discrimination is to consider a coding system in which *the first spike from multiple sources is used to transmit information between stages of processing.*

In relation to the proposed model of static form processing, these data suggest the visual system can operate to resolve the appearance of unexpected objects primarily in a feed-forward manner, without the need for lateral inhibition or feed-back loops: a property few models embody. This feed-forward processing does not deny one possible role of top-down processing suggested above.

Processing efficiency of parallel distributed processing models

Predictions based on consideration of the functioning and architecture of parallel distributed processing models were assessed using a simple simulation of an interactive activation competition model. Unlike previous model simulations where noise was introduced into the model input, here the noise was introduced within the processing elements themselves. This type of noise allowed comparison between the performance of network models (typically with noiseless processing elements) and neural signals (with noisy signals).

It was confirmed that the predictions of slow rise time and slowly emergent discrimination between stimuli was correct for the interactive activation competition model tested. It was argued that these properties are present in all models using feedback and lateral inhibitory connections. Such properties lie in marked contrast to the performance of STPa cells sensitive to different views of the head. The slowly emergent discrimination did not reflect a change in the input-output function of the artificial network. Indeed, the delay in discrimination closely followed that predicted from simple statistical considerations. The network did also show sensitivity to the level of noise, in that at higher levels (> 5%), the model became unstable and that unpredictable pseudo-steady states were reached, even in the absence of any external input. The levels of noise used to test the IAC model were substantially less than those seen in the population and Average Cell analyses for STPa cells.

These results indicate form processing in the macaque monkey operates in a feed-forward way with remarkable efficiency compared to many artificial

models. If computational models are to aid the unravelling of the primate visual system, it would appear that there needs to be a shift in emphasis from the use of lateral inhibition and feedback to a more "feed-forward" system. The brain's use of multiple cortical areas for visual form processing may allow for this primarily feed-forward style of processing.

2. MOTION PROCESSING

Directional motion processing in area STPa

An investigation was made into the directional sensitivity of cells in the macaque area STPa to the motion of objects. The cells studied were sensitive to the presence of motion but showed little or no selectivity for the form of the stimulus.

Directional tuning was not continuously distributed about all possible directions. The majority of cells were most responsive to motion in a direction within 15 degrees of one of the three cartesian axes (up/down, left/right, towards/away). Tuning to direction varied in sharpness. For most (34/37) cells the angular change in direction required to reduce response to half maximal was between 45 and 70^{o} (for 3/37 cells it was $> 90^{o}$).

The estimates of the directionality (median Id = 0.97) of STPa cells was similar to that reported for posterior motion processing areas (MT, the middle temporal area and MST, the medial superior temporal area). The tuning for direction (sharpness, distribution and discrimination) of the motion sensitive STPa cells were found to be similar to the tuning for perspective view of STPa cells selective for static form of the head and body (Perrett et al. 1991).

The responses of motion sensitive STPa cells occurred at an earlier latency (mean 91 ms) than responses of cells selective for static form (mean 119 ms) but the time course of responses of the two classes of cell were similar in many other respects. On average the STPa responses showed a 100-300 ms transient burst of

activity followed by a tonic discharge maintained at approximately 20% of the peak firing rate for the duration of stimulation. The early response latency and directional selectivity indicate that motion sensitivity in STPa cells derives from the dorsal visual pathway via MT/MST. The similarity of tuning for direction and perspective view within STPa may facilitate the integration of motion and form processing within this high level brain area.

Form-from-motion processing in STPa

Investigation of the extent to which cells in STPa are responsive under 'biological motion' conditions where the form of the body is defined only by the movement of light patches attached to the points of limb articulation was made.

One third of the cells (25/72) selective for the form and motion of walking bodies, showed sensitivity to the moving light displays. 7 of these cells showed only partial sensitivity to form from motion, in so far as the cells responded more to moving light displays than to moving controls but failed to discriminate body view. These 7 cells exhibited directional selectivity.

18 cells showed statistical discrimination for both direction of movement and body view under biological motion conditions. Most of the 18 cells showed reduced responses to the impoverished moving light stimuli compared to full light conditions. The 18 cells were thus sensitive to detailed form information (body view) from the pattern of articulating motion. Cellular processing of the global pattern of articulation was indicated by the observations that none of the cells were found sensitive to movement of individual limbs and that jumbling the pattern of moving limbs reduced response magnitude.

A further 10 cells were tested for sensitivity to moving light displays of whole body actions other than walking. 5/10 of these cells showed selectivity for form displayed by biological motion stimuli that paralleled the selectivity under normal lighting conditions. The cell responses thus provide direct evidence for neural mechanisms computing form from non-rigid motion.

Overall the study indicated that the selectivity of the cells was for body view, specific direction and specific type of body motion presented by moving light displays and is not predicted by many current computational approaches to the extraction of form from motion.

3. CONJOINT FORM AND MOTION PROCESSING

Integration of form and motion

Processing of visual information is known to occur in at least 2 separate cortical pathways, commonly labelled the "What" and "Where" or the "Form" and "Motion" pathways. This division lies in marked contrast to our everyday experience in which we have a unified percept of both the form and motion of objects, implying integration of both types of information.

A neuronal population in area STPa was investigated which were selectively responsive to the sight of bodies moving through the environment (e.g. left profile view of the body moving to the viewer's left). A total of 161 cells were found to be sensitive to the body form and motion.

Ten of these cells 161 (6%) showed selectivity for form and motion in that good responses were obtained for the sight of any body view (but not control objects) moving in any direction. For 13 other cells (8%), maximal responses were only obtained when a particular body view was moving in any direction. The majority of the remaining cells (125/138, 91%) responded to only one combination of view and direction (termed unimodal cells, e.g. left profile view moving left, not right profile moving left or left profile moving right). A smaller number of cells (9) responded selectively to two, opposite, combinations of view and direction (termed bimodal cells, e.g. left profile moving left and right profile moving right but not other view and direction combinations). A minority of cells (4) were found to show "object-centered" selectivity to view and direction

combinations, responding to all directions of motion where the body moves in a direction compatible with the direction it faces. Some cells responded to left profile go left, right profile go right, face view moving towards, back view moving away but not other view and direction combinations and other cells responded to the body moving incompatibly (e.g. responding to left profile go right, right profile go left, face view moving away, back view moving towards but not other view and direction combinations).

Approximately three quarters of the neurons (106/138, 77%) showed selectivity for compatible motion (e.g. left profile move left), and one quarter showed selectivity for incompatible motion (e.g. right profile moving left).

It was shown that the response of some of these cells was selective for the motion and form of the same object not simply the juxtaposition of appropriate form and motion signals.

Mechanisms of Integration

The cells selective for conjoint form and motion information in area STPa were compared to two other cell populations in the same brain area: (a) those sensitive to static presentation of stimulus form (see chapter 3), and (b) those sensitive to direction of motion but not stimulus form (see chapter 5). These latter two cell populations share characteristics which would facilitate combination of their outputs to produce cells sensitive to conjoint information about form and its motion (similar tuning functions and preferential coding of particular directions and body views).

The response latencies of cells selective for form and motion are on average coincident with cells selective for direction of motion (but not stimulus form). Both these cell populations, however, have earlier response latencies than cells selective for static form.

Cells in STPa responsive to conjoint information about form and motion were assessed for the relative strengths of isolated motion and form inputs. A

small number of the cells (4%) responsive to preferred form when the stimuli were static showed response suppression to the inappropriate direction of motion.

For the cells conjointly selective to form and motion the majority (95%) of the responses were characterized as showing non-linear summation of form and motion inputs.

The degree of directional discrimination of cells conjointly sensitive to form and motion was measured. Cells sensitive to conjoint form and motion showed significantly reduced directional discrimination compared with cells sensitive to direction alone. The degree of view discrimination of cells conjointly sensitive to form and motion was also assessed. Cells sensitive to conjoint form and motion showed a small but significantly reduced view discrimination compared with cells sensitive to form independent of motion.

A simple model was proposed which combined outputs from cells sensitive to motion alone with outputs of cells sensitive to form independent of motion to generate cells selective for conjoint form and motion stimuli. This model closely predicts the observed changes in the direction and view discrimination between these three cell populations.

Temporal aspects of Integration

Analyses were performed to investigate the timing for discrimination of form information and motion information in the responses of cells conjointly selective for form and motion in area STPa of the macaque monkey.

Calculation of the average of Early response latency cells (cells whose response latency was under the sample mean) suggests that direction information is present in cell responses some 35 ms before form information becomes evident. Direction and form information become evident within 5 ms of each other in the average Late response latency cells (those cells whose response latency was greater than the sample mean).

Inputs relating to movement show an initial response period which does not discriminate direction. The quality of initial direction discrimination appeared to be independent of response latency. The initial discrimination of form was related to response latency such that cells with longer response latencies showed greater initial discrimination of form in their responses. It is argued that this correlation reflects the observation that form inputs to area STPa arrive some 20 ms after motion inputs into area STPa.

Temporal considerations of the response and discrimination profiles suggested that neither oscillations nor synchrony are used to enable the integration of form and motion information about the same object.

4. POSSIBLE FURTHER APPLICATIONS OF TIME COURSE ANALYSIS

Much of the neurobiological data presented in this thesis was analysed looking at short (5 ms) time slices. While it is important to realize that neurons generally operate over longer time periods, the use of this analysis technique allowed constraints to be placed on computational approaches to vision (in particular chapters 3 and 9). Below, a few other applications for such analysis are considered.

Response latency and generalization

STPa cells response show a degree of generalization over object components (Wachsmuth et al. 1994) and response tolerance to changes in orientation, size and perspective view. In chapter 1, it was hypothesized that capacity to generalize will increase with response latency (see also Perrett and Oram 1993; Oram and Perrett 1994). The details of any such relationship could be investigated using techniques developed here and would be informative. Constancies for different attributes (orientation, size, view point) could emerge at the same time and collectively (i.e. cells will show similar levels of constancies

within each dimension, with the extent of generalization in all dimensions increasing with latency of response). This would imply similar mechanisms for each type of constancy. Alternatively constancies may be established independently and in parallel (the degree of constancy for one dimension would be independent of the others). Generalizations could also emerge sequentially with invariance across position occurring first, then across size, lighting, view and finally across component.

The timing of extra-retinal effects

A recent article by Desimone and colleagues (Chellazi et al. 1993) showed clear effects of distractor targets on infero-temporal cell responses during a match to sample task. In the article, they noted that there was considerable delay (about 200 ms) between response onset and the effect of the distractors becoming apparent.

Investigation of any relationship between the time at which such effects become apparent and the nature of the task would provide insights into the possible source or sources of the modulation. In particular, it would be interesting to correlate the latency of the attention effects with the number of distractors present in the match display.

A paradigm similar to that of Chelazzi et al. (1993) could also be used to investigate high level neural representations that might underlie "pop-out" effects. Neural responses in area V1 have been postulated as underlying some "pop-out" effects: analysis of IT cell responses could show a marked difference in either the temporal or magnitude domain of the responses when the visual search was for a "pop-out" target versus a "non pop-out" target.

Learning of associations

The work of Miyashita and colleagues (Miyashita et al., in press, Nature) has suggested that the learning of associations between input patterns (fractals)

involves the hippocampal complex (in particular the entorhinal cortex). They also present evidence that suggests that the maintenance of these associations also involves the same cortical areas.

A plausible hypothesis that explains the results of Miyashita et al. (in press) is that the learned associations between paired associates operate via a simple loop: IT selective cells selective for one stimulus of a pair feeding onto the hippocampal complex and the feedback from the hippocampal complex acting as IT input of the second, "associated" pattern. This could be directly tested by comparing when the information (discrimination) of each of the pair of patterns became evident within the response.

In summary the techniques developed in this thesis used to describe the time course of discrimination of inputs by cell responses could be used to aid understanding of the neural basis of several psychological visual recognition phenomena.

REFERENCES

Adelson, E.H., & Movshon, J.A. (1982). Phenomenal coherence of moving visual plaids. *Nature, 300, 523-525.*

Aggleton, J.P. (1993). The contribution of the amygdala to normal and abnormal emotional states. *Trends in the Neurosciences, 16, 328-333.*

Aggleton, J.P., Butron, M.J., & Passingham, R.E. (1980). Cortical and subcortical afferents to the amygdala of the rhesus monkey (*Macaca mulatta*). *Brain Research, 190, 347-368.*

Albrecht, D.G., & Geisler, W.S. (1994). Visual cortex neurons in monkey and cat: Contrast response nonlinearities and stimulus selectivity. In (T.Lawton Ed) *Computational Vision Based on Neurobiology, Proceedings SPIE, 2054, 12-31.*

Albright, T.D. (1984). Direction and orientation selectivity of neurons in the visual area MT of the macaque. *Journal of Neurophysiology, 52, 1106-1130.*

Albright, T.D. (1992). Form-cue invarient motion processing in primate visual cortex. *Science, 255, 1141-1143.*

Albright, T.D., Desimone, R., & Gross, C.G. (1984). Columnar organization of directionally selective cells in visual area MT of the macaque. *Journal of Neurophysiology, 51, 16-31.*

Allison, T., Ginter, H., McCarthy, G., Nobre, A.C., Puce, A., Luby, M., & Spencer, D.D. (1994). Face recognition in human extrastriate cortex. *Journal of Neurophysiology, 71, 821-825.*

Allman, J., Miezin, F. & McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local-global comparisons in visual neurons. *Annual Review of Neuroscience, 8, 407-430.*

Alonso, J.M., Cudeiro, J., Perez, R., Gonzalez, F., & Acuna, C. (1993). Orientation influences of layer V of visual area 18 upon cells in layer V of area 17 in the cat. *Experimental Brain Research, 96, 212-220.*

Amaral, D.G., Price, J.L., Pitkanen, A., & Carmichael, S.T. (1992). Anatomical organization of the primate amygdaloid complex. In (J.P. Aggleton, Ed) *The Amygdala*, pp 1-66, Wiley-Liss, New York, U.S.A.

Andersen, R.A., & Siegel R.M. (1989). Motion processing in the primate cortex. In (G.M. Edelman, W.E. Gall & W.M. Cowan, Eds) *Signal and Sense: Local and Global Order in Perceptual Maps*, pp 163-184.

Baddely, R. (1993). Spatial correlation and the geometry of striate cortex. *Perception, 22, Suppl 69a.*

Baizer, J.S, Desimone, R., & Ungerleider, L.G. (1993). Comparison of subcortical connections of inferior temporal and posterior parietal cortex in monkeys. *Visual Neuroscience, 10, 59-72.*

Baizer, J.S., Ungerleider, L.G., & Desimone, R. (1991). Organisation of visual inputs to the inferior temporal and posterior parietal cortex in macaques. *Journal of Neuroscience, 11, 168-190.*

Barlow, H., & Foldiak, P. (1989). Adaptation and decorrelation in the cortex. In (R. Durbin, C. Miall & G. Mitchison, Eds), *The Computing Neuron*, pp 54-72). Addison-Wesley, Wokingham, England.

Barlow, H.B. (1972). Single units and sensation: A neural doctrine for perceptual psychology. *Perception, 1, 371-394.*

Barlow, H.B. (1985). The twelfth Bartlett Memorial Lecture: The role of single neurons in the psychology of perception. *Quarterly Journal of Experimental Psychology, 37A, 121-145.*

Barlow, H.B. (1994). The neuron doctrine in perception. In (M. Gazanigga ed) *The Cognitive Neurosciences* (Boston, MIT Press) pp 415-435

Barlow, H.B., & Levick, W.R. (1965). The mechanism of directionally selective units in rabbit's retina. *Journal of Physiology (London), 178, 477-504.*

Bassili, J.N. (1979). Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality and Social Psychology, 37, 2049-2058.*

Baylis, G.C., Rolls, E.T., & Leonard, C.M. (1985). Selectivity between faces in the responses of a population of neurons in the cortex of the superior temporal sulcus of the macaque monkey. *Brain Research, 342, 91-102.*

Baylis, G.C., Rolls, E.T., & Leonard, C.M. (1987). Functional subdivisions of the temporal lobe neocortex. *Journal of Neuroscience, 7, 330-342.*

Benevento, L.A., & Yoshida, K. (1981). The afferent and efferent organization of the lateral geniculo-prestriate pathways in the macaque monkey. *Journal of Comparative Neurology, 203, 455-474.*

Bennett, B.M., & Hoffman, D.D. (1985). The computation of structure from fixed-axis motion: Nonrigid structures. *Biological Cybernetics, 51, 293-300.*

Benson, P.J, & Perrett, D.I. (1991). Perception and recognition of photographic quality facial caricatures: Implications for the recognition of natural images. *European Journal of Cognitive Psychology, 3, 105-135.*

Berman, N.J., Douglas, R.J., Martin, K.A.C., & Whitteridge, D. (1991). Mechanisms of inhibition in cat visual cortex. *Journal of Physiology (London), 440, 697-722.*

Bertenthal, B.I., Proffitt, D.R., Spenter, N.B., & Thomas, M.A. (1985). The development of infant sensitivity to biomechanical motions. *Child Developement, 56, 531-543.*

Bertrand, M. (1969). The behavioural repertoire of the stumptail macaque. *Bibliotheca Primatologica*, 11. Karger, Basel.

Best, J., Mallot, H., Kruger, K., & Dinse, H.R.O. (1989). Dynamics of visual information processing in cortical systems. In (L. Personnaz & G. Dreyfus Eds), *Neural Networks: From Models to Applications*, pp 107-116. I.D.S.E.T., Paris, France.

Best, J., Reuss, S., & Dinse, H.R.O. (1986). Lamina-specific differences of visual latencies following photic stimulation in the cat striate cortex. *Brain Research, 385, 356-360.*

Biederman, I. (1987). Recognition by components: A theory of human image understanding. *Psychological Review, 94, 115-145.*

Blasdel, G.G. (1992). Differential imaging of ocular dominance and orientation selectivity in monkey striate cortex. *Journal of Neuroscience, 12, 3115-3138.*

Blasdel, G.G. (1992). Orientation selectivity, preference, and continuity in monkey striate cortex. *Journal of Neuroscience, 12, 3139-3181.*

Blum, B. (1985). Manipulation reach and visual reach neurons in the inferior parietal lobule of the rhesus monkey. *Behavioural Brian research, 18, 167-173.*

Blum, B. (1989). Inferior parietal lobule neurons related to eye and hand reach movements in rhesus monkey. *Brain and Behaviour Evolution, 33, 104-108.*

Boussaoud, D., Ungerleider L.G., & Desimone R. (1990). Pathways for motion analysis: Cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. *Journal of Comparative Neurology, 296, 462-495.*

Britten, K.H., Newsome, W.T., & Saunders, R.C. (1992). Effects of inferotemporal cortex lesions on form-from-motion discrimination in monkeys. *Experimental Brain Research, 88, 292-302.*

Brothers, L., & Ring, B. (1992). A neuroethological framework for the representation of minds. *Journal of Cognitive Neuroscience, 4, 107-118.*

Brothers, L., & Ring, B. (1993). Mesial temporal neurons in the macaque monkey with responses selective for aspects of social stimuli. *Behavioural Brain Research, 57, 53-61.*

Bruce, C.J., Desimone, R., & Gross, C.G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology, 46, 369-384.*

Bruce, C.J., Desimone, R., & Gross, C.G. (1986). Both striate cortex and superior colliculus contribute to visual properties of neurons in superior temporal polysensory area of macaque monkey. *Journal of Neurophysiology, 55, 1057-1075.*

Bullier, J., & Kennedy, H. (1983). Projection of the lateral geniculate nucleus onto cortical area V2 in the macaque monkey. *Experimental Brain Research, 53, 168-172.*

Burton, A.M, Bruce, V., & Johnston, R.A. (1990). Understanding face recognition with an interactive activation model. *British Journal of Psychology, 81, 361-380.*

Cahusac, P.M.B., Rolls, E.T., & Marriott, F.H.C. (1991). Potentiation of neuronal responses to natural visual input paired with postsynaptic activation in the hippocampus of the awake monkey. *Neuroscience Letters, 124, 39-43.*

Campbell, F.W., & Robson J.G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology (London), 197, 551-566.*

Campbell, R. (1992). The neuropsychology of lip-reading. *Philosophical Transaction of the Royal Society of London (Series B), 335, 39-45.*

Chapman, B., Zahs, K.R., & Stryker, M.P. (1991). Relation of cortical cell orientation selectivity to alignment of receptive fields of the geniculocortical afferents that arborize within a single orientation column in ferret visual cortex. *Journal of Neuroscience, 11, 1347-1358.*

Chelazzi, L., Miller, E.K., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature, 363, 345-347.*

Clothiaux, E.E., Bear, M.F., & Cooper L.N. (1991). Synaptic plasticity in visual cortex: Comparison of theory with experiment. *Journal of Neurophysiology, 66, 1785-1804.*

Coburn, K.L., Ashford, J.W., & Fuster, J.M. (1990). Visual response latencies in temporal lobe structures as a function of stimulus information load. *Behavioural Neuroscience, 104, 62-73.*

Colby, C.L. (1991) The neuroanatomy and neurophysiology of attention. *Journal of Child Neurology, 6, S90-S118.*

Colby, C.L., & Miller, E.K. (1986) Eye movement related responses of neurons in superior temporal polysensory area of macaque. *Society of Neuroscience Abstracts, 12, 1.*

Colby, C.L., Duhamel, J.R., & Goldberg, M.E. (1993). Ventral intraparietal area of the macaque: Anatomical location and visual response properties. *Journal of Neurophysiology, 69, 902-914.*

Corbetta, M., Miezin, F.M., Dobmeyer, S., Shulman, G.L. & Peteresen S.E. (1990). Attentional modulation of neural processing of shape, colour and velocity in humans. *Science, 248, 1556-1559.*

Cowey, A., & Macar, V.L. (1992). The effect of removing superior temporal cortical motion areas in the macaque monkey: I. Motion discrimination using simple dots. *European Journal of Neuroscience, 4, 1219-1227.*

Cutting, J.E. (1978). Generation of synthetic male and female walkers through manipulation of a biomechanical invariant. *Perception, 7, 393-405.*

Cutting, J.E., & Kozlowski L.T. (1977). Recognising friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomics Society, 9, 353-356.*

Cutting, J.E., Moore, C., & Morrison, R. (1988). Masking the motion of human gait. *Perception and Psychophysics, 44, 339-347.*

Cutting, J.E., Proffitt, D.R., & Kozlowski, L.T. (1978). A biomechanical invariant for gait perception. *Journal of Experimental Psychology, 4, 357-372.*

De Valois, R.L., Albrecth, D.G., & Thorell, L.G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research, 22, 545-559.*

De Valois, R.L., Yund, E.W., & Hepler, N. (1982). The orientation and direction selectivity of cells in macaque visual cortex. *Vision Research, 22, 531-544.*

De Yoe, F.A., & Van-Essen, D.C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neuroscience, 11, 219-226.*

Dean, P. (1976). Effects of inferotemporal lesions on the behaviour of monkeys. *Psychological Bulletin, 83, 41-71.*

Desimone, R., Albright, T.D., Gross, C.G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience, 8, 2051-2062.*

Desimone, R., Wessinger, M., Thomas, L., & Schneider, W. (1990). Attentional control of visual perception, Cortical and subcortical mechanisms. In *Cold Spring Harbor Symposia on Quantitative Biology,* Vol. 55, pp 963-971. Cold Spring Harbor Laboratory Press, U.S.A.

Dittrich, W.D. (1993). Action categories and the perception of biological motion. *Perception, 22, 15-22.*

Douglas, R.J. and Martin, K.A.C. (1991). A functional microcircuit for cat visual cortex. *Journal of Physiology (London), 440, 735-769.*

Douglas, R.J., Martin, K.A.C., & Whitteridge, D. (1991). An intracellular analysis of the visual responses of neurones in cat visual cortex. *Journal of Physiology (London), 440, 659-696.*

Dubner, R., & Zeki, S.M. (1971). Response properties and recptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain Research, 35, 528-532.*

Duffy, C.J., & Wurtz, R.H. (1991a) Sensitivity of MST neurons to optic flow stimuli. I: A continuum of response selectivity to large-field stimuli. *Journal of Neurophysiology, 65, 1329-1345.*

Duffy, C.J., & Wurtz, R.H. (1991b). Sensitivity of MST neurons to optic flow stimuli. II: Mechanisms of response selectivity revealed by small-field stimuli. *Journal of Neurophysiology, 65, 1346-1359.*

Duhamel, J.R., Colby, C.L., & Goldberg, M.E. (1992). The updating of the representation of visual space in parietal cortex by intended eye-movements. *Science, 255, 90-92.*

Edelman, G.M. (1978). Group selection and phasic re-entrant signaling: A theory of higher brain function. In (V.B. Mountcastle & G.M. Edelman Eds), *The Mindful Brain*, pp 51-100. MIT Press, Cambridge, MA, U.S.A.

Ekman, P., & Friesen, W.V. (1976). *Pictures of facial affect.* Consulting Psychologists Press, Palo Alto, U.S.A.

Engel, A.K., Konig, P., & Schillen, T.B. (1992). Why does the cortex oscillate?. *Current Biology, 2, 332-334.*

Engel, A.K., Konig, P., Kreiter, A.K., Schillen, T.B., & Singer, W. (1992). Temporal coding in the visual cortex: new vistas on integration in the nervous system. *Trends in Neuroscience, 15, 218-226.*

Erickson, R.G., & Thier, P. (1991). A neuronal correlate of spatial stability during periods of self-induced visual motion. *Experimental Brain Research, 86, 608-616.*

Eskandar, E.N., Optican, L.M., & Richmond, B.J. (1992b). Role of inferior temporal neurons in visual memory: II Multiplying temporal waveforms related to vision and memory. *Journal of Neurophysiology, 68, 1296-1306.*

Eskandar, E.N., Richmond, B.J., & Optican, L.M. (1992a). Role of inferior temporal neurons in visual memory: I Temporal encoding of information about visual images, recalled images, and behavioural context. *Journal of Neurophysiology, 68, 1277-1295.*

Fahy, F.L., Riches, I.P. & Brown, M.W. (1993). Neuronal activity related to visual recognition memory: Long-term memory and the encoding of recency and familiarity information in the primate anterior and medial inferior temporal and rhinal cortex. *Experimental Brain Research, 96, 457-472.*

Faugier-Grimaud, S., Frenois, C., & Stein, D.G. (1978). Effects of posterior parietal lesions on visually guided behaviour in monkeys. *Neuropsychologia, 16, 151-168.*

Felleman, D.J., & Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1, 1-47.*

Ferrera, V.P., Nealey, T.A., & Maunsell, J.H.R. (1992). Mixed parvocellular and magnocellular geniculate signals in visual area V4. *Nature, 358, 756-758.*

Ferrera, V.P., Nealey, T.A., & Maunsell, J.H.R. (1994). Responses in macaque visual area V4 following inactivation of the parvocellular and magnocellular LNG pathways. *Journal of Neuroscience, 14, 2080-2088.*

Field D.J. (1994). What is the goal of sensory coding? *Neural Computation, 6, 559-601.*

Finlay, B.L., Schiller, P.H., & Volman, S.F. (1976). Quantitative studies of single cell properties in monkey striate cortex IV: Corticotectal cells. *Journal of Neurophysiology, 39, 1352-1361.*

Fiorani, M., Rosa, M.G.P., Gattass, R., & Rocha-Miranda, C.E. (1992). Dynamic surrounds of receptive fields in primate striate cortex: A physiological

basis for perception completion? *Proceedings of the National Academy of Science (U.S.A.), 89, 8547-8551.*

Foldiak, P. (1989). Adaptive network for optimal linear feature extraction. *Proceedings of the IEEE/INNS International Joint Conference on Neural Networks, Washington DC, 401-405.*

Foldiak, P. (1990). Forming sparse representation by local anti-Hebbian learning. *Biological Cybernetics, 64, 165-170.*

Foldiak, P. (1991). Learning invariance from transformation sequences. *Neural Computation, 3, 194-200.*

Foldiak, P. (1993). The 'ideal homunculus': Statistical inference from neural population responses. In (F.H Eeckman & J.M. Bower, Eds), *Computation and Neural System*, pp 55-60. Kluwer Academic Publishers, Norwell, MA, USA.

Fox, R., & McDaniel, C. (1982). The perception of biological motion by human infants. *Science, 218, 486-487.*

Fries, W. (1981). The projection from the lateral geniculate nucleus to the prestriate cortex of the macaque monkey. *Proceedings of the Royal Society of London (Series B), 213, 73-80.*

Frohn, H., Geiger, H., & Singer, W. (1987). A self-organizing neural network sharing features of the mammalian visual system. *Biological Cybernetics, 55, 333-343.*

Fujita, I., Tanaka, K., Ito, M., & Cheng, K. (1992). Columns for visual features of objects in monkey inferotemporal cortex. *Nature, 360, 343-346.*

Fukushima, K. (1980). Neocognition: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics, 36, 193-202.*

Fuster, J.M. (1990). Inferotemporal units in selective visual attention and short-term memory. *Journal of Neurophysiology, 64, 681-697.*

Gabor, D. (1946). Theory of communication. *Journal of the Institute of Electrical Engineers (London), 93, 429-457.*

Gaffan, D., & Harrison, S. (1993). Role of the dorsal prestriate cortex in visuospatial configural discrimination by monkeys. *Behaviour and Brain Research, 56, 119-125.*

Gallant, J.L., Braun, J., & Van Essen, D.C. (1993). Selectivity for polar, hyperbolic and cartesian gratings in macaque visual cortex. *Science, 259, 100-103.*

Galletti, C., Battaglini, P.P., & Aicari, G. (1988). 'Real-motion' cells in visual area V2 of behaving macaque monkeys. *Experimental Brain Research, 69, 279-288.*

Galletti, C., Battaglini, P.P., & Fattori, P. (1990). 'Real-motion' cells in area V3A of macaque visual cortex. *Experimental Brain Research, 82, 67-76.*

Galletti, C., Squatrito, S, Battaglini, P.P., & Maioli, M.G. (1984). 'Real-motion' cells in the primary visual cortex of macaque monkeys. *Brain Research, 301, 95-110.*

Gattas, R., & Gross, C.G. (1981). Visual topography of striate projection zone in posterior superior temporal sulcus of the macaque. *Journal of Neurophysiology, 46, 621-638.*

Gawne, T.J., & Richmond, B.J. (1993). How independent are the messages carried by adjacent inferior temporal cortical neurons? *Journal of Neuroscience, 13, 2758-2771.*

Gawne, T.J., McClurkin, J.W., Richmond, B.J., & Optican, L.M. (1991b). Lateral geniculate neurons in behaving primates III. Response predictions of a channel model with multiple spatial-to-temporal filters. *Journal of Neurophysiology, 66, 809-823.*

Gawne, T.J., Richmond, B.J., & Opticam, L.M. (1991a). Interactive effects among several stimulus parameters on the responses of striate cortical complex cells. *Journal of Neurophysiology, 66, 379-389.*

Geisler, W.S., Albrecht, D.G., Salvi, R.J., & Saunders, S.S. (1991). Discrimination performance of single neurons: Rate and temporal-pattern information. *Journal of Neurophysiology, 66, 334-362.*

Ghose, G.M., & Freeman, R.D. (1992). Oscillatory discharge in the visual system: Does it have a functional role? *Journal of Neurophysiology, 68, 1558-1574.*

Girard, P., & Bullier, J. (1989). Visual activity in area V2 during reversible inactivation of area 17 in the macaque monkey. *Journal of Neurophysiology, 62, 1287-1302.*

Girard, P., Salin, P.A., & Bullier, J. (1991). Visual activity in areas V3a and V3 during reversible inactivation of area V1 in the macaque monkey. *Journal of Neurophysiology, 66, 1493-1503.*

Gochin, P.M., Miller, E.K., Gross, C.G., & Gerstein, G.L. (1991). Functional interactions among neurons in inferior temporal cortex of the awake macaque. *Experimental Brain Research, 84, 505-516.*

Goddard, N.H. (1992). The perception of articulated motion: Recognizing moving light displays. PhD thesis, University of Rochester, New York, U.S.A.

Goodale, M.A. (1993). Visual pathways supporting perception and action in the primate cerebral cortex. *Current Biology, 3, 578-585.*

Goodale, M.A., & Milner, A.D. (1992). Seperate visual pathways for perception and action. *Trends in Neuroscience, 15, 20-25.*

Goodale, M.A., Jakobson, L.S., Milner, A.D., Perrett, D.I., Benoson, P.J., & Hietanen, J.K. (1994). The nature and limits of orientation and pattern procesing supporting visuomotor control in a visual form agnosic. *Journal of Cognitive Neuroscience, 6, 46-56.*

Grasse, K.L., & Cynader, M.S. (1982). Electrophysiology of medial terminal nucleus of accessory optic system in the cat. *Journal of Neurophysiology, 48, 490-504.*

Grasse, K.L., & Cynader, M.S. (1984). Electrophysiology of lateral and dorsal terminal nuclei of the cat accessory optic system. *Journal of Neurophysiology, 51, 276-293.*

Grasse, K.L., Cynader, M.S., & Douglas, R.M. (1984). Alterations in response properties in the lateral and dorsal terminal nuclei of the cat accessory optic system following visual cortex lesions. *Experimental Brain Research, 55, 69-80.*

Gray, C.M., Konig, P., Engel, A.K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnal synchronization which reflects global stimulus properties. *Nature, 338, 334-337.*

Grosof, D.H., Shapley, R.M., & Hawken, M.J. (1993). Macaque V1 neurons can signal 'illusory' contours. *Nature, 365, 550-552.*

Gross, C.G. (1991). Contribution of striate cortex and the superior colliculus to visual function in area MT, the superior temporal polysensory area and inferior temporal cortex. *Neuropsychologia, 29, 497-515.*

Gross, C.G. (1992). Representation of visual stimuli in inferior temporal cortex. *Philosophical Transactions of the Royal Society of London (Series B), 335, 3-10.*

Gross, C.G., Rocha-Miranda, C.E., & Bender, D.B. (1972). Visual properties of neurons in inferotemporal cortex of the monkey. *Journal of Neurophysiology, 35, 96-111.*

Grossberg, S. (1987). Competitive learning: From interactive activation to adaptive resonance. *Cognitive Science, 11, 23-63.*

Harries, M.H., & Perrett, D.I. (1991). Modular organization of face processing in temporal cortex: Physiological evidence and possible anatomical correlates. *Journal of Cognitive Neuroscience, 3, 9-24.*

Hasselmo, M.E., Rolls, E.T., & Baylis, G.C. (1989). The role of expression and identity in the face selective response of neurons in the temporal visual cortex of the monkey. *Behavioural Brain Research, 32, 203-218.*

Hasselmo, M.E., Rolls, E.T., Baylis, G.C., & Nalwa, V. (1989). Object-centered encoding by face-selective neurons in the cortex in the superior temporal sulcus of the monkey. *Experimental Brain Research, 75, 417-429.*

Hata, Y., Tsumoto, T., Sato, H., & Tamura, H. (1991). Horizontal interactions between visual cortical neurones studied by cross-correlation analysis in the cat. *Journal of Physiology (London), 441, 593-614.*

Hawken, M.J., & Parker, A.J. (1987). Spatial properties of neurons in the monkey striate cortex. *Proceedings of the Royal Society of London (Series B), 231, 251-288.*

Heeley, D.W., & Buchanan-Smith, H.M. (1992). Directionality acuity for drifting plaids. *Vision Research, 32, 97-104.*

Heywood, C.A. & Cowey, A. (1987). On the role of cortical area V4 in the discrimination of hue and pattern in macaque monkeys. *Journal of Neuroscience, 7, 2601-2617.*

Heywood, C.A., & Cowey, A. (1992). Cortical area V4 and its role in the perception of color. *Journal of Neuroscience, 12, 4056-4065.*

Heywood, C.A., Shields, C., & Cowey, A. (1988). The involvement of the temporal lobes in colour discrimination. *Experimental Brain Research, 71, 437-441.*

Hietanen, J.K. (1992). Visual processing in a primate temporal association cortex: Insensitivity to self-induced motion. Unpublished thesis, St. Andrews University.

Hietanen, J.K., & Perrett, D.I. (1993). Motion sensitive cells in the macaque superior temporal polysensory area: I. lack of response to the sight of the monkey's own hand. *Experimental Brain Research, 93, 117-128.*

Hietanen, J.K., Perrett, D.I., Oram, M.W., Benson, P.J., & Dittrich, W.H. (1992). The effects of lighting conditions on responses of cells selective for face views in the macaque temporal cortex. *Experimental Brain Research, 89, 157-171.*

Hikosaka, K., Iwai, E., Saito, H-A., & Tanaka, K. (1988). Polysensory properties of neurons in the anterior bank of the caudal superior temporal sulcus of the macaque monkey. *Journal of Neurophysiology, 60, 1615-1637.*

Hildreth, E.C., & Koch, C. (1987). The analysis of visual motion: From computational theory to neuronal mechanisms. *Annual Review of Neuroscience, 10, 477-533.*

Hirano, T. (1990). Synaptic transmission between rat inferior olivary neurons and cerebellar Purkinje cells in culture. *Journal of Neurophysiology, 63, 181-189.*

Hirsch, J.A. and Gilbert C.D. (1991). Synaptic physiology of horizontal connections in the cat's visual cortex. *Journal of Neuroscience, 11, 1800-1809.*

Hoffman, D.D., & Flinchbaugh, B.E. (1982). The interpretation of biological motion. *Biological Cybernetics, 42, 195-204.*

Hoffman, K.P., & Distler, C. (1989). Quantitative analysis of visual receptive fields of neurons in nucleus of the optic tract and dorsal terminal nucleus of the accessory optic tract in macaque monkey. *Journal of Neurophysiology, 62, 416-428.*

Hubel, D.H., & Livingstone M.S. (1987). Segregation of form, color and stereopsis in primate area 18. *Journal of Neuroscience, 7, 3378-.*

Hubel, D.H., & Livingstone M.S. (1990). Color and contrast sensitivity in the lateral geniculate body and primary visual cortex of the macaque monkey. *Journal of Neuroscience, 10, 2223-2237.*

Hubel, D.H., & Wiesel, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology (London), 195, 215-243.*

Humphreys, G.W., Donnelly, N., & Riddoch, M.J. (1993) Expression is computed separately from facial identity and it is computed separately for moving and static faces: Neuropsychological evidence. *Neuropsychologia, 31, 173-181.*

Humphreys, G.W., Riddoch M.J, Quinlan P.T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology, 1, 67-103.*

Iwai, E., Yukie, M., Umistu, Y., Kido, S., & Niihara, T. (1980). Geniculo-prestriate projection in the macaque monkey. *Experimental Brain Research, 41 Supplement, A19.*

Jansson, G., & Johansson, G. (1973). Visual perception of bending motion. *Perception, 2, 321-326.*

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics, 14, 201-211.*

Johansson, G. (1976). Spatio-temporal differentiation and integration in visual motion perception. *Psychological Research, 38, 379-393.*

Kendrick, K.M., & Baldwin, B.A. (1987). Cells in temporal cortex of conscious sheep can respond preferentially to the sight of faces. *Science, 236, 448-450.*

Knierim, J.J., & Van-Essen, D.C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology, 67, 961-980.*

Kobatake, E. & Tanaka, K. (1992) Selective neuronal responses to complex visual object-features are already present in posterior part of the macaque inferotemporal cortex. *Society of Neuroscience Abstracts, 17, 443.*

Kobatake, E. & Tanaka, K. (1993) Selectivity for features beyond orientation, color, size and simple texture in the prestirate areas V2 and V4. *Society of Neuroscience Abstracts 18, 146.*

Kobatake, E., & Tanaka, K. (1994). Neural selectivities to complex object-features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology, 71, 856-867.*

Koenderink, J.J., & van Doorn, A.J. (1979). The internal representation of solid shape with respect to vision. *Biological Cybernetics, 32, 211-216.*

Komatsu, H., & Ideura, Y. (1993). Relationships between color, shape and pattern selectivities of Neurons in the inferior temporal cortex of the Monkey. *Journal of Neurophysiology, 70, 677-694.*

Komatsu, H., & Wurtz, R.H. (1988a). Relation of cortical ares MT and MST to pursuit eye movements. I. Localization and visual response properties of neurons. *Journal of Neurophysiology, 60, 580-603.*

Komatsu, H., & Wurtz, R.H. (1988b). Relation of cortical areas MT and MST to pursuit eye movements. III. Interaction with full-field visual stimulation. *Journal of Neurophysiology, 60, 621-644.*

Komatsu, H., & Wurtz, R.H. (1989). Modulation of pursuit eye movements by stimulation of cortical areas MT and MST. *Journal of Neurophysiology, 62, 31-47.*

Konorski, J. (1967). *Integrative Activity of the Brain* Chicago, Illinois, University of Chicago Press.

Kozlowski, L.T., & Cutting J.E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception and Psychophysics, 21, 575-580.*

Kreiter, A.K., & Singer, W. (1992). Oscillatory neuronal responses in the visual cortex of the awake macaque monkey. *European Journal of Neuroscience, 4, 369-375.*

Krone, G., Mallot, H.A., Palm, G., & Schuz, A. (1986). Spatio-temporal fields: A dynamical model derived from cortical archiotectonics. *Proceedings of the Royal Society of London (Series B), 226, 421-444.*

Kruger, J., & Becker, J.D. (1991). Recognizing the visual stimulus from neuronal discharges. *Trends in Neuroscience, 14, 282-286.*

Lamme, V.A.F., van Dijk B.W., & Spekreijse H. (1994). Organization of contour from motion processing in primate visual cortex. *Vision Research, 34, 721-735.*

Lappin, J.S., Doner, J.F., & Kottas, B. (1980). Minimal conditions for the visual detection of structure and motion in three dimensions. *Science, 209, 717-719.*

Lee, H.J., & Chen, Z. (1985). Determination of 3D human body postures from a single view. *Computer Vision, Graphics and Image Processing, 30, 148-168.*

Lehky, S.R., Sejnowski, T.J., & Desimone, R. (1992). Predicting responses of nonlinear neurons in monkey striate cortex to complex patterns. *Journal of Neuroscience, 12, 3568-3581.*

Leung, M.K., & Yang, Y.H. (1987). A region-based approach for human body motion analysis. *Pattern Recognition, 20, 321-339.*

Levinson, E., & Sekuler, R. (1980). A two-dimensional analysis of direction-specific adaptation. *Vision Research, 20, 103-107.*

Li, L., Miller, E.K., & Desimone, R. (1993). The representation of stimulus familiarity in anterior inferior temporal cortex. *Journal of Neurophysiology, 69, 1918-1929.*

Lowe, D.G. (1987). *Perceptual organization and visual recognition.* Kluwer Academic Publishers, Boston U.S.A.

Lund, J.S. (1988). Anatomical organization of macaque monkey striate visual cortex. *Annual Review of Neuroscience, 11, 253-288.*

Macar, V.L., & Cowey, A. (1992). The effect of removing superior temporal cortical motion areas in the macaque monkey. II Motion discrimination

using random dot displays. *European Journal of Neuroscience, 4, 1228-1238.*

Mackay, W.A., Blum, B., & Mendonca, A.J. (1992). Visual responses to reward-related cues in inferior parietal lobule. *Ophthal. Physiol. Opt, 12, 209-214.*

Mansfield, R.J.W., & Ronner, S.F. (1978). Orientation anisotropy in monkey visual cortex. *Brain Research, 149, 229-234.*

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* Freeman, San Francisco, CA, U.S.A..

Marr, D., & Nishihara, H.K. (1978). Representation and recognition of the spatial organization of three dimensional shapes. *Proceedings of the Royal Society of London (Series B), 200, 269-294.*

Marr, D., & Vaina, L. (1982). Representation and recognition of the movements of shapes. *Proceedings of the Royal Society of London (Series B), 214, 501-524.*

Martin (1992) Parallel pathways converge. *Current Biology, 2, 555-557.*

Mason, A., Nicoll, A., & Stratford, K. (1991). Synaptic transmission between individual pyramidal neurons of the rat visual cortex in vitro. *Journal of Neuroscience, 11, 72-84.*

Mather, G., Radford, K., & West, S. (1992). Low level visual processing of biological motion. *Proceedings of the Royal Society of London (Series B), 249, 149-155.*

Maunsell, J.H.R. (1994). Commentry in (T.Lawton Ed) *Computational Vision Based on Neurobiology, Proceedings SPIE, 2054.*

Maunsell, J.H.R., & Gibson, J.R. (1992). Visual response latencies in striate cortex of the macaque monkey. *Journal of Neurophysiology, 68, 1332-1344.*

Maunsell, J.H.R., & Newsome, W.T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience, 10, 363-401.*

Maunsell, J.H.R., & Van Essen, D.C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey I: Selectivity for stimulus direction speed and orientation. *Journal of Neurophysiology, 49, 1127-1147.*

Maunsell, J.H.R., Nealey, T.A., & De Priest, D.D. (1990). Magnocellular and parvocellular contributions to responses in the middle temporal visual area (MT) of the macaque monkey. *Journal of Neuroscience, 10, 3323-3334.*

McClelland, J.L. (1991). Stochastic interactive processes and the effect of context on perception. *Cognitive Psychology, 23, 1-44.*

McClelland, J.L., & Rumelhart, D.E. (1988). Interactive activation and competition. In *Explorations in Parallel Distributed Processing: A Handbook of Models, Programs and Exercises*, pp 11-47. MIT Press, Cambridge, MA, U.S.A..

McClelland, J.L., Rumelhart, D.E., & Hinton, G.E. (1986). The appeal of parallel distributed processing. In (D.E. Rumelhart, J.L. McClelland & the PDP Research Group, Eds), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1, pp 3-44. MIT Press, Cambridge, MA, U.S.A..

McClurkin, J.W., Gawne, T.J., Optican, L.M., & Richmond, B.J. (1991b). Lateral geniculate neurons in behaving primates II. Encoding of visual information in the temporal shape of the response. *Journal of Neurophysiology, 66, 794-808.*

McClurkin, J.W., Gawne, T.J., Richmond, B.J., Optican, L.M., & Robinson, D.L. (1991a). Lateral geniculate neurons in behaving primates I. Responses to two-dimensional stimuli. *Journal of Neurophysiology, 66, 777-793.*

Merigan, W.H., & Maunsell, J.H.R. (1993). How parallel are the primate visual pathways? *Annual Review of Neuroscience, 16, 369-402.*

Merigan, W.H., Byrne, C.E., & Maunsell, J.H.R. (1991). Does primate motion perception depend on the magnocellular pathway? *Journal of Neuroscience, 11, 3422-3429.*

Michael, C.R. (1981). Columnar organization of color cells in monkey's striate cortex. *Journal of Neurophysiology, 46, 587-604.*

Mikami, A., Newsome, W.T., & Wurtz, R.H. (1986a). Motion sensitivity in macaque visual cortex. I: Mechanisms of direction and speed selectivity in extrastriate area MT. *Journal of Neurophysiology, 55, 1308-1327.*

Mikami, A., Newsome, W.T., & Wurtz, R.H. (1986b). Motion selectivity in macaque visual cortex. II: Spatiotemporal range of directional interactions in MT and V1. *Journal of Neurophysiology, 55, 1328-1339.*

Miles, R. (1990). Synaptic excitation of inhibitory cells by single CA3 hippocampal pyramidal cells of the guinea-pig in vitro. *Journal of Physiology (London), 428, 61-77.*

Miles, R., & Wong, R.K.S. (1986). Excitatory synaptic interactions between CA3 neurons in guinea-pig hippocampus. *Journal of Physiology (London), 373, 397-418.*

Miller, E.K., Li, L., & Desimone, R. (1993). Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *Journal of Neuroscience, 13, 1460-1478.*

Milner, A.D., & Goodale, M.A. (1993). Visual pathways to perception and action. *Progress in Brain Research, 95, 317-337.*

Milner, A.D., Ockleford, E.M., & Dewer, W. (1977). Visuo-spatial performance following posterior parietal and lateral frontal lesions in stumptail macaques. *Cortex, 13, 350-360.*

Milner, A.D., Perrett, D.I., Johnston, R., Benson, P.J., Jordon, T.R., Heeley, D.W., Bettucci, D., Mortara, F., Mutani, R., Terazzi, E. & Davidson, D.L.W. (1991). Perception and action in `visual form agnosia'. *Brain, 114, 405-428.*

Minsky, and Papert. 1969.

Mishkin, M., Ungerleider, L.G., & Macko, K.A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in the Neuroscience, 6, 414-417.*

Mistlin, A.J., & Perrett, D.I. (1990). Visual and somatosensory processing in the macaque temporal cortex: The role of 'expectation'. *Experimental Brain Research, 82, 437-450.*

Miyashita, Y. (1990). Associative representation of the visual long-term memory in the neurons of the primate temporal cortex. In (E. Iwai & M. Mishkin Eds), *Vision, Memory and the Temporal Lobe*, pp 75-87. Elsevier Science Publishing, New York, U.S.A..

Miyashita, Y., & Chang, H-S. (1988). Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature, 331, 68-70.*

Miyashita, Y., Date, A., & Okuno H. (1993). Configurational encoding of complex visual forms by single neurons of monkey temporal cortex. *Neuropsychologia, 31, 1119-1131.*

Miyashita, Y., Sakai, K., & Okuno, H. (1993). Long-term memory neurons related to tuning and association mechanisms in primate vision. *22nd Congress of the international Union of Physiological Sciences, p 43, 27.2.*

Mizuno, N., Uchida, K., Nomura, S., Nakamura, Y., Sugimoto, T., & Uemura-Sumi, M. (1981). Extrageniculate projections to the visual cortex in the macaque monkey: an HRP study. *Brain Research, 212, 454-459.*

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science, 229, 782-784.*

Motter, B.C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *Journal of Neurophysiology, 70, 909-919.*

Movshon, J.A., Adelson, E.H., Gizzi, M.S., & Newsome, W.T. (1985). The analysis of moving visual patterns. In (C. Chagas, R. Gattass & C. Gross, Eds), *Pattern Recognition Mechanisms,* pp 117-151. Springer-Verlag, New York, U.S.A.

Muly, E.C., & Fitzpatrick, D. (1992). The morphological basis for binocular and ON/OFF convergence in tree shrew striate cortex. *Journal of Neuroscience, 12, 1319-1334.*

Murasugi, C.M., Salzman, C.D., & Newsome, W.T. (1993). Microstimulation in visual area MT: Effects of varying pulse amplitude and frequency. *Journal of Neuroscience, 13, 1719-1729.*

Newcombe, F., & Russell, W.R. (1969). Dissociated visual perceptual and spatial deficits in focal lesions of the right hemisphere. *Journal of Neurology, Neurosurgery and Psychiatry, 32, 73-81.*

Newsome, W.T., & Pare, E.B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *Journal of Neuroscience, 8, 2201-2211.*

Newsome, W.T., Britten, K.H., Salzman, C.D., & Movshon, J.A. (1990). Neuronal mechanisms of motion perception. In *Cold Spring Harbor Symposia on Quantitative Biology*, 55, 697-705. Cold Spring Harbor Press, U.S.A.

Newsome, W.T., Mikami, A., & Wurtz, R.H. (1986). Motion selectivity in macaque visual cortex. III: Psychophysics and physiology of apparent motion. *Journal of Neurophysiology, 55, 1340-1351.*

Newsome, W.T., Wurtz, R.H., & Komatsu, H. (1988). Relation of cortical areas MT and MST to pursuit eye movements: II Differentiation of retinal to extraretinal inputs. *Journal of Neurophysiology, 60, 604-620.*

Ono, T., Eifuku, S., Nakamura K., & Nishijo, H. (1991). Monkey hippocampal neuron responses related to spatial and non-spatial influence. *Neuroscience Letters, 159, 75-78.*

Optican, L.M., & Richmond, B.J. (1987). Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex III: Information theoretic analysis. *Journal of Neurophysiology, 57, 162-178.*

Oram M.W. & Perrett D.I. (1994). Responses of anterior superior temporal polysensory (STPa) neurons to `biological motion' stimuli. *Journal of Cognitive Neuroscience, 6, 99-116.*

Oram M.W., Perrett, D.I., & Hietanen, J.K. (1993). Directional tuning of motion-sensitive cells in the anterior superior temporal polysensory area of the macaque. *Experimental Brain Research, 97, 274-294.*

Oram, M.W., & Perrett, D.I. (1992). Time course of neural responses discriminating different views of the face and head. *Journal of Neurophysiology, 68, 70-84.*

Oram, M.W., & Perrett, D.I. (1994). Modelling visual recognition from neurobiological constraints. *Neural Networks, in press.*

Oram, M.W., & Perrett, D.I. (1994). Neural processing of biological motion in the macaque temporal cortex. In (T.Lawton, Ed) *Computational Vision based on Neurobiology, Proceedings of SPIE, 2054, 155-165.*

Pece, A.E.C. (1993). Modelling and simulations of the feedback loop between visual cortex and LGN. *Perception, 22 Supplement, 42c.*

Perrett, D.I., & Oram, M.W. (1993). The neurophysiology of shape processing. *Image and Vision Computing, 11, 317-333.*

Perrett, D.I., Harries, M.H., Benson, P.J., Chitty, A.J., Mistlin, A.J. (1990). Retrieval of structure from rigid and biological motion: an analysis of the visual response of neurons in the macaque temporal cortex. In (A. Blake & T. Troscianko, Eds), *AI and the Eye*, pp 181-201. John Wiley, Chichester, England.

Perrett, D.I., Harries, M.H., Bevan, R., Thomas, S., Benson, P.J., Mistlin, A.J., Chitty, A.J., Hietanen, J.K., & Ortega, J.E. (1989). Frameworks of

analysis for the neural representation of animate objects and actions. *Journal of Experimental Biology, 146, 87-114.*

Perrett, D.I., Harries, M.H., Chitty, A.J., & Mistlin, A.J. (1990). Three stages in the classification of body movements by visual neurons. In (H.B. Barlow, C. Blakemore & M. Weston-Smith, Eds), *Images and Understanding*, pp 94-107. Cambridge University Press, Cambridge, England.

Perrett, D.I., Hietanen, J.K., Oram, M.W., & Benson, P.J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London (Series B), 335, 23-30.*

Perrett, D.I., Mistlin, A.J., Chitty, A.J., Smith P.A.J. Potter, D.D., Broennimann, R., & Harries, M.H. (1988). Specialised face processing and hemispheric asymmetry in man and monkey: Evidence from single unit and reaction time studies. *Behavioural Brain Research, 29, 245-258.*

Perrett, D.I., Mistlin, A.J., Harries, M.H., & Chitty, A.J. (1989). Understanding the visual appearance and consequence of hand actions. In *Vision and Action: The Control of Grasping*, pp 163-180. Ablex Publishing, Norwood, New Jersey, U.S.A..

Perrett, D.I., Oram, M.W., & Wachsmuth, E. (1993). Understanding minds and expression from facial signals: Studies at the brain cell level. *Proceedings of the 2nd IEEE International Workshop on Robot and Human Communication Supplement, 3-5 November 1993, Tokyo Japan*, p 3-12.

Perrett, D.I., Oram, M.W., Harries, M.H., Bevan, R., Hietanen, J.K., Benson, P.J., & Thomas, S. (1991). Viewer-centred and object-centred coding of heads in the macaque temporal cortex. *Experimental Brain Research, 86, 159-173.*

Perrett, D.I., Rolls, E.T., & Caan, W. (1982). Visual neurons responsive to faces in the monkey temporal cortex. *Experimental Brain Research, 47, 329-342.*

Perrett, D.I., Smith, P.A.J., Mistlin, A.J., Chitty, A.J., Head, A.S., Potter, D.D., Broennimann, R., Milner, A.D., & Jeeves, M.A. (1985a). Visual analysis of body movements by neurons in the temporal cortex of the macaque monkey: a preliminary report. *Behavioural Brain Research, 16, 153-170.*

Perrett, D.I., Smith, P.A.J., Potter, D.D., Mistlin, A.J., Head, A.S., Milner, A.D., & Jeeves, M.A. (1985b). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London (Series B), 223, 293-317.*

Perrett, D.I., Smith, P.A.J., Potter, D.D., Mistlin, A.J., Head, A.S., Milner, A.D., & Jeeves, M.A. (1984). Neurons responsive to faces in the temporal cortex: studies of functional organization sensitivity to identity and relation to perception. *Human Neurobiology, 3, 197-208.*

Peterhans, E., & von der Heydt, R. (1989a). Mechanisms of contour perception in monkey visual cortex: Contours bridging gaps. *Journal of Neuroscience, 9, 1749-1763.*

Peterhans, E., & von der Heydt, R. (1989b). Subjective contours - bridging the gap between psychophysics and physiology. *Trends in Neuroscience, 14, 112-118.*

Peterhans, E., von der Heydt, R., & Baumgartner, G. (1986). Neuronal responses to illusory contour stimuli reveal stages of visual cortical processing. In (J.D. Pettigrew, K.J. Sanderson & W.R. Levick, Eds), *Visual Neuroscience*, pp 343-351. Cambridge University Press, Cambridge, England.

Peterhans, P., & von der Heydt, R. (1993). Functional organization of area V2 in the alert macaque. *European Journal of Neuroscience, 5, 509-524.*

Poizner, H., Bellugi, U., & Lutes-Driscol, V. (1981). Perception of American Sign Language in dynamic point-light displays. *Journal of Experimental Psychology, Human Perception and Performance, 7, 430-440.*

Probst, Th., Plendl, H., Paulus, W., Wist, E.R., & Scherg, M. (1993). Identification of the visual motion area (area V5) in the human brain by dipole source analysis. *Experimental Brain Research, 93, 345-351.*

Proffitt, D.R., Bertenthal, B.I. & Roberts Jr, R.J. (1984) The role of occlusion in reducing multistability in moving point-light displays. *Perception and Psychophysics, 36, 315-323.*

Rashid, R.F. (1980). Towards a system for the interpretation of moving light displays. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-2, 574-581.*

Reid, R.C., Soodak, R.E., & Shapley, R.M. (1991). Directional selectivity and spatiotemporal structure of receptive fields of simple cells in cat striate cortex. *Journal of Neurophysiology, 66, 505-529.*

Rhodes, G., Brennan, S.E., & Carey, S. (1987). Identification and Ratings of Caricatures: Implications for Mental Representations of Faces. *Cognitive Psychology, 19, 473-497.*

Riches, I.P., Wilson, F.A.W., & Brown, M.W. (1991). The effects of visual stimulation and memory on neurons of the hippocampal formation and the neighboring parahippocampal gyrus and inferior temporal cortex of the primate. *Journal of Neuroscience, 11, 1763-1779.*

Richmond, B.J., & Optican, L.M. (1987). Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex II: Quantification of response waveform. *Journal of Neurophysiology, 57, 147-161.*

Richmond, B.J., & Sato, T. (1987). Enhancement of inferior temporal neurons during visual discrimination. *Journal of Neurophysiology, 58, 1292-1306.*

Richmond, B.J., Optican, L.M., Podell, M., & Spitzer, H. (1987). Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex I: Response characteristics. *Journal of Neurophysiology, 57, 132-146.*

Richmond, B.J., Wurtz, R.H., & Sato, T. (1983). Visual response of inferior temporal neurons in awake rhesus monkey. *Journal of Neurophysiology, 50, 1415-1432.*

Ridley, R.M., Hester, N.S., & Ettlinger, G. (1977). Stimulus- and response-dependent units from the occipital and temporal lobes of the unanaesthetised monkey performing learnt visual tasks. *Experimental Brain Research, 27, 539-552.*

Robinson, D.L., & Rugg, M.D. (1988). Latencies of visually responsive neurons in various regions of the rhesus monkey brain and their relation to human visual responses. *Biological Psychology, 26, 111-116.*

Rodman, H.R., & Albright, T.D. (1989). Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Experimental Brain Research, 75, 53-64.*

Rodman, H.R., Gross, C.G., & Albright, T.D. (1989). Afferent basis of visual response properties in area MT of the macaque.I: Effects of striate cortex removal. *Journal of Neuroscience, 9, 2033-2050.*

Rodman, H.R., Gross, C.G., & Albright, T.D. (1990). Afferent basis of visual response properties in area MT of th macaque. II. Effects of superior colliculus removal. *Journal of Neuroscience, 10, 1154-1164.*

Rolls, E.T. (1984). Neurons in the cortex of the temporal lobe and amygdala of the monkey with responses selective for faces. *Human Neurobiology, 3, 209-222.*

Rolls, E.T. (1992). Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. *Philosophical Transactions of the Royal Society of London Series B, 335, 11-22.*

Rolls, E.T., & Baylis, C.G. (1986). Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the macaque monkey. *Experimental Brain Research, 65, 38-48.*

Rolls, E.T., Baylis, G.C., & Hasselmo, M.E. (1987). The responses of neurons in the cortex in the superior temporal sulcus of the monkey to band-pass spatial frequency filtered faces. *Vision Research, 27, 311-326.*

Rolls, E.T., Baylis, G.C., & Leonard, C.M. (1985). Role of low and high spatial frequencies in the face-selective responses of neurons in the cortex in the superior temporal sulcus in the monkey. *Vision Research, 25, 1021-1035.*

Rolls, E.T., Baylis, G.C., Hasselmo, M.E., & Nalwa, V. (1989). The effect of learning on the face selective properties of neurons in the cortex in the superior temporal sulcus of the monkey. *Experimental Brain Research, 76, 153-164.*

Rolls, E.T., Perrett, D.I., Caan, W., & Wilson, F. (1982). Neuronal responses related to visual recognition. *Brain, 105, 611-646.*

Rosene, D.L., Roy, N.J., Davis, B.J. (1986). A cryoprotection method that facilitates cutting frozen sections of whole monkey brains for histological and histochemical processing without freezing artifact. *Journal of Histochemistry Cytochemistry, 34, 1301-1315.*

Roy, J-P, & Wurtz, R.H. (1990). The role of disparity-sensitive cortical neurons in signalling the direction of self-motion. *Nature, 348, 160-162.*

Roy, J-P., Komatsu, H., & Wurtz, R.H. (1992). Disparity sensitivity of neurons in monkey extrastriate area MST. *Journal of Neuroscience, 12, 2476-2492.*

Rumelhart, D.E., & Zipser, D. (1986). Feature Discovery by Competitive Learning. In (D.E. Rumelhart, J.L. McClelland & the PDP Research Group, Eds), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition,* Vol. 1, pp 151-193. MIT Press, Cambridge, MA, U.S.A..

Rybak, I.A., Shevtsova, N.A., Podladchikova, L.N., & Golovan, A.V. (1991). A visual cortex domain model and its use for visual information processing. *Neural Networks, 4, 3-13.*

Saito, H., Tanaka, K., Isono, H., Yasuda, M., & Mikami, A. (1989). Directionally selective response of cells in the middle temporal area (MT) of the macaque monkey to the movement of equiluminous opponent color stimuli. *Experimental Brain Research, 75, 1-14.*

Saito, H., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y., & Iwai, E. (1986). Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *Journal of Neuroscience, 6, 145-157.*

Sakai, K. and Miyashita, Y. (1991). Neural organization for the long-term memory of paired associates. *Nature, 354, 152-155.*

Sakata, H. (1991). Neural correlates of the perception of depth rotation in the parietal association cortex of the monkey. *Biomedical Research, 2, 149-151.*

Sakata, H., & Kusunoki, M. (1992). Organization of space perception: Neural representation of three-dimensional space in the posterior parietal cortex. *Current Opinion in Neurobiology, 2, 170-174.*

Salzman, C.D., Britten, K.H., & Newsome, W.T. (1990). Cortical microstimulation influences judgements of motion detection. *Nature, 346, 174-177.*

Salzman, C.D., Murasugi, C.M., Britten, K.H., & Newsome, W.T. (1992). Microstimulation in visual area MT: Effects on direction discrimination performance. *Journal of Neuroscience, 12, 2331-2355.*

Sary, G., Vogels, R., & Orban, G.A. (1993). Cue invariant shape selectivity of macaque inferior temporal neurons. *Science, 260, 995-997.*

Sato, T. (1988). Effects of attention and stimulus interaction on visual responses on inferior temporal neurons in macaque. *Journal of Neurophysiology, 60, 344-364.*

Sato, T. (1989). Interaction of visual stimuli in the receptive fields of inferior temporal neurons in awake macaques. *Experimental Brain Research, 77, 23-30.*

Sato, T., Kawamura, T., & Iwai, E. (1980). Responsiveness of inferotemporal single units to visual pattern stimuli in monkeys performing discrimination. *Experimental Brain Research, 38, 313-319.*

Sayer, R.J., Friedlander, M.J., & Redman, S.J. (1990). The time course and amplitude of EPSPs evoked at synapses between pairs of CA3/CA1 neurons in the hippocampal slice. *Journal of Neuroscience, 10, 826-836.*

Schiller, P.H. (1993). The effects of V4 and middle temporal (MT) area lesions on visual performance in the rhesus monkey. *Visual Neuroscience, 10, 717-746.*

Schiller, P.H., Finlay, B.L., & Volman, S.F. (1976). Quantitative studies of single-cell properties in monkey striate cortex I: Spatiotemporal organization of receptive fields. *Journal of Neurophysiology, 39, 1288-1319.*

Schiller, P.H., Finlay, B.L., & Volman, S.F. (1976). Quantitative studies of single-cell properties in monkey striate cortex III: Spatial frequency. *Journal of Neurophysiology, 39, 1334-1351.*

Schiller, P.H., Finlay, B.L., & Volman, S.F. (1976). Quantitative studies of single cell properties in monkey striate cortex II: Orientation specificity and ocular dominance. *Journal of Neurophysiology, 39, 1320-1333.*

Schwartz, E.L. (1980). Computational anatomy and functional architecture of the striate cortex: A spatial mapping approach to perceptual coding. *Vision Research, 20, 645-.*

Schwartz, E.L., Desimone, R., Albright, T.D., & Gross, C.G. (1983). Shape recognition and inferior temporal neurons. *Proceedings of the National Academy of Science (U.S.A.), 80, 5776-5778.*

Seibert, M., & Waxman, A.M. (1989). Spreading activation layers, visual saccades, and invariant representation for neural pattern recognition systems. *Neural Networks, 2, 9-27.*

Seibert, M., & Waxman, A.M. (1991) Learning aspect graph representations from view sequences. In D S Touretzky (Ed), *Advances in Neural Network Information Processing Systems*, Vol. 2, pp 258-265. Morgan Kaufman, San Mateo, CA, U.S.A..

Seibert, M., & Waxman, A.M. (1992) Learning and recognizing 3D obects from multiple views in a neural system. In (H. Wechsler Ed), *Neural Networks for Perception*, Vol 1 (2), pp 426-444. Academic Press, U.S.A..

Seibert, M., & Waxman, A.M. (1992). Adaptive 3D object recognition from multiple views. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 14, 107-124.*

Seibert, M., & Waxman, A.M. (1993). An approach to face recognition using saliency maps and caricatures. In *Proceedings World Congress on Neural Networks*, Vol. 3, pp 661-664, Portland Oregan, U.S.A.

Seltzer, B., & Pandya, D.N. (1978). Afferent cortical connections and architectonics of the superior temporal sulcus and surrounding cortex in the Rhesus monkey. *Brain Research, 149, 1-24.*

Sergent, J., & Signoret, J.L. (1992a). Varieties of functional deficits in prospagnosia. *Cerebral Cortex, 2, 375-388.*

Sergent, J., & Signoret, J.L. (1992b). Implicit access to knowledge derived from unrecognized faces in prospagnosia. *Cerebral Cortex, 2, 389-400.*

Shipp, S. & Zeki, S.M. (1985). Segregation of pathways leading from area V2 to areas V4 and V5 of macaque monkey visual cortex. *Nature, 315, 322.*

Siegel, R.M., & Andersen, R.A. (1986). Motion perceptual deficits following ibotenic acid lesions of the middle temporal area in the behaving rhesus monkey. *Society of Neuroscience Abstracts, 12, 1183.*

Siegel, R.M., & Andersen, R.A. (1990). The perception of structure from visual motion in monkey and man. *Journal of Cognitive Neuroscience, 2, 306-319.*

Sillito, A.M. (1975). The contribution of inhibitory mechanisms to the receptive field properties of neurones in the striate cortex of the cat. *Journal of Physiology (London), 250, 305-329.*

Sillito, A.M. (1975). The effectiveness of biculline as an antagonist of GABA and visually evoked inhibition in the cat's striate cortex. *Journal of Physiology (London), 250, 287-304.*

Simpson, J.I., Leonard, C.S., & Soodak, R.E. (1988). The accessory optic system of rabbit. II. Spatial organization of direction selectivity. *Journal of Neurophysiology, 60, 2055-2072.*

Singer, W. (1990). Self-organization of cognitive structures. In (J.C. Eccels & O. Creutzfeldt, Eds), *The principle of design and operation of the brain*, pp 119-135. Springer-Verlag, Berlin, Germany.

Singer, W. (1993). Neuronal representations, assemblies and temporal coherence. In (T.P. Hicks, S. Molotchnikoff and T. Ono Eds) *Progress in Brain Research, 95, 461-473.*

Singer, W., Gray, C., Engel, A., Konig, P., Artola, A., & Brocher, S. (1990). Formation of cortical cell assemblies. In *Cold Spring Harbour Symposia on Quantitative Biology,* Vol. 55, pp 939-952. Cold Spring Harbour Laboratory Press, U.S.A..

Snedecor, G.W., & Cochran, G.W. (1980). *Statistical Methods.* Iowa State University Press, Ames, Iowa, U.S.A..

Snowden, R.J., Treue, S., & Andersen, R.A. (1992). The response of neurons in areas V1 and MT of the alert rhesus monkey to moving random dot patterns. *Experimental Brain Research, 88, 389-400.*

Snowden, R.J., Treue, S., Erickson, R.G., & Andersen, R.A. (1991). The response of area MT and V1 neurons to transparent motion. *Journal of Neuroscience, 11, 2768-2785.*

Soodak, R.E., & Simpson, J.I. (1988). The accessory optic system of rabbit. I. Basic visual response properties. *Journal of Neurophysiology, 60, 2037-2054.*

Squire, L.R. (1992). Memory and the hippocampus: A sythesis from findings with rats, monkeys, and humans. *Psychological Review, 99, 195-231.*

Squire, L.R., Shimamura, A.P., & Amaral D.G. (1989). Memory and the hippocampus. In (J.H. Byrne & W.D. Berry, Eds), *Neural Models of Plasticity*, pp 208-239. Academic Press, San Diego, U.S.A..

Squire, L.R., Zola-Morgan, S., Cave, C.B., Haist F., Musen G., & Suzuki W.A. (1990). Memory: Organization of brain systems and cognition. In *Cold Sping Harbour Symposia on Quantitative Biology*, Vol. 55, pp 1007-1023. Cold Spring Harbour Laboratory Press, U.S.A..

Stark, D.G., & Wilson H.R. (1990). Do Gabor functions provide appropriate descriptions of visual cortical receptive fields? *Journal of Opthamological Society of America (Part A), 7, 1362-1373.*

Stoner, G.R., & Albright, T.D. (1993). Image segmentation cues in motion processing: Implications for modularity in vision. *Journal of Cognitive Neuroscience, 5, 129-149.*

Sugie, N., & Kato, K. (1987). A computational model of biological motion perception. *IEEE Montech, Conference on Biomedical Technologies, November 1987, 140-143.*

Sugihara, K., & Sugie, N. (1984). Recovery of rigid structure from orthographically projected optical-flow. *Computer Vision, Graphics and Image Processing, 27, 309-320.*

Sugita, Y. & Tanaka, K. (1991). Occlusion-related cue used for analysis of motion in the primate visual cortex. *NeuroReport, 2, 751-754.*

Sumi, S. (1984). Upside-down presentation of the Johansson moving light-spot pattern. *Perception, 13, 283-286.*

Swindale, N.V. (1990). Is the cerebral cortex modular? *Trends in Neuroscience, 13, 487-492.*

Tanaka, K. (1993). Neuronal mechanisms of object recognition. *Science, 262, 685-688.*

Tanaka, K., & Saito, H-A. (1989). Analysis of motion of the visual field by direction, expansion/contraction and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology, 62, 626-641.*

Tanaka, K., Fujita, I., Kobatake, E., Cheng, K., & Ito, M. (1993). Serial processing of visual object-features in the posterior and anterior parts of the inferotemporal cortex. In (T. Ono, L.R. Squire, M.E. Raichle, D.I. Perrett & M. Fukuda, Eds), *Brain Mechanisms of Perception and Memory: From Neuron to Behaviour*, pp 34-46. Oxford University Press, Oxford, England.

Tanaka, K., Fukada, Y., & Saito, H. (1989). Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology, 62, 642-656.*

Tanaka, K., Hikosaka, K., Saito, H-A., Yukie, M., Fukada, Y., & Iwai, E. (1986). Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *Journal of Neuroscience, 6, 134-144.*

Tanaka, K., Saito, H., Fukada, Y., & Moriya, M. (1991). Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *Journal of Neurophysiology, 66, 170-189.*

Thier, P., & Erickson, R.G. (1992). Responses of visual-tracking neurons from cortical area MST-l to visual, eye and head motion. *European Journal of Neuroscience, 4, 539-553.*

Thorpe, S.J. (1990). Spike arrival times: A highly efficient coding scheme for neural networks. In (R. Eckmiller, G. Hartmann & G. Hauske, Eds),

Parallel Processing in Neural Systems and Computers, pp 91-94. Elsevier, Amsterdam, The Netherlands.

Thorpe, S.J., & Imbert, M. (1989). Biological constraints on connectionist models. In (R. Pfeifer, Z. Schreter and F. Fogelman-Soulié, Eds) *Connectionism in Perspective*, pp 613-622. Elsevier, Amsterdam, The Netherlands.

Thorpe, S.J., Celebrini, S., Trotter, Y., Pouget, A., & Imbert M. (1989). Dynamic aspects of orientation coding in V1 of the awake primate. *European Journal of Neuroscience, Supplement, 322.*

Torre, V., & Poggio, T. (1978). A synaptic mechanism possibly underlying directional selectivity to motion. *Proceedings Royal Society London (Series B), 202, 409-416.*

Tovee, M.J., & Rolls, E.T. (1992). Oscillatory activity is not evident in the primate temporal visual cortex with static stimuli. *NeuroReport, 3, 369-372.*

Tovee, M.J., & Rolls, E.T. (1993). The temporal responses of translation-invariant neurons in the primate temporal visual cortex. *Perception, 22 Supplement, 141.*

Tovee, M.J., Rolls, E.T., Treves, A., & Bellis, R.P. (1993). Information endcoding and the responses of single neurons in the primate temporal visual cortex. *Journal of Neurophysiology, 70, 640-654.*

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12, 97-136.*

Treisman, A., & Paterson, R. (1984). Emergent features, attention, and object perception. *Journal of Experimental Psychology: Human Perception and Performance, 10, 12-31.*

Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance, 16, 459-478.*

Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology, 14, 107-141.*

Trotter, Y., Thorpe, S.J., Pouget, A., Celebrini and S. Imbert, M. (1989). Processing of orientation in V1 of the awake monkey. *Society of Neuroscience Abstracts, 15, 1056.*

Turner, B.H., Mishkin, M., & Knapp M. (1980). Organization of the amygdalopetal projections from modality-specific cortical association areas in the monkey. *Journal of Comparative Neruology, 191, 515-543.*

Ullman, S. (1979). The interpretation of structure from motion. *Proceedings of the Royal Society of London (Series B), 203, 405-426.*

Ungerleider, L.G., & Mishkin, M. (1982) Two cortical visual systems. In (D.J. Ingle, M.A. Goodale & R.J.W. Mansfield, Eds) *Analysis of visual behaviour*, pp 549-586. MIT Press, Cambridge, MA, U.S.A..

Vaina, L.M., Lemay, M., Bienfang, D.C., Choi, A.Y., & Nakayama, K. (1990). Intact "biological motion" and "structure from motion" perception in a patient with impaired motion mechanisms: A case study. *Visual Neuroscience, 5, 353-369.*

Van Essen, D.C. (1993). What's in a receptive field? Filters vs feature detectors revisited. *Brain Research Association Abstracts, 10, 17.*

Van Essen, D.C., & Zeki, S. (1978). The topographic organization of rhesus monkey presriate cortex. *Journal of Physiology (London), 277, 193-226.*

Van Essen, D.C., Maunsell, J.H.R., & Bixby, J.L. (1981). The middle temporal visual area in the macaque: myeloarchitecture, connections, functional properties and topographic organization. *Journal of Comparitive Neurology, 199, 293-326.*

Vogels, R., & Orban, G.A. (1991). Quantitative study of striate single unit responses in monkey performing an orientation task. *Experimental Brain Research, 84, 1-11.*

von der Heydt, R., & Peterhans, E. (1989). Mechanisms of contour perception in monkey visual cortex: II Lines of pattern discontinuity. *Journal of Neuroscience, 9, 1731-1748.*

von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science, 224, 1260-1262.*

von der Heydt, R., Peterhans, E., & Dursteler, M.R. (1992). Periodic-pattern-selective cells in monkey visual cortex. *Journal of Neuroscience, 12, 1416-1434.*

von der Marlsburg, C. (1988). Pattern recognition by labelled graph matching. *Neural Networks, 1, 141-148.*

von der Marlsburg, C., & Schneider, W. (1986). A neural cocktail-party processor. *Biological Cybernetics, 54, 29-40.*

von Seelen, W., Mallot, H.A., & Giannakopoulos, F. (1987). Characteristics of neuronal systems in the visual cortex. *Biological Cybernetics, 56, 37-49.*

Wachsmuth, E., Oram, M.W., & Perrett, D.I. (1993). Response of single units to objects and their components in the anterior superior temporal polysensory area (STPa) of the macaque. *Perception, 22 Supplement, 138c.*

Wachsmuth, E., Oram, M.W., & Perrett, D.I. (1994). Recognition of objects and their component parts: Responses of single units in the temporal cortex of the macaque. In press, *Cerebral Cortex.*

Walmsley, B., & Stuklis, R. (1989). Effects of spatial and temporal dispersion of synaptic input on the time course of synaptic potentials. *Journal of Neurophysiology, 61, 681-687.*

Walsh, V., Butler, S.R., Carden, D., & Kulikowski, J.J. (1992a). The effects of V4 lesions on the visual abilities of macaques: Shape discrimination. *Behavioural Brain Research, 50, 115-126.*

Walsh, V., Kulikowski, J.J., Butler, S.R., & Carden, D. (1992b). The effects of V4 lesions on the visual abilities of macaques: Colour categorization. *Behavioral Brain Research, 52, 81-89.*

Warrington, E.K. (1986). Visual deficits associated with occipital lesions in man. *Experimental Brain Research, Suplementum Series, 11, 247-261.*

Watson, J.D.G., Meyers, R., Frackowiak, R.S.J., Hajnal, J.V., Woods, R.P., Mazziotta, J.C., Shipp, S. & Zeki, S. (1993). Area V5 of the human brain: Evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral Cortex, 3, 79-94.*

Webb, J., & Aggarwal, J. (1982). Structure from motion of rigid and jointed objects. *Artificial Intelligence, 19, 107-130.*

Weiskrantz, L., & Saunders, S. (1984). Impairments of visual object transforms in monkeys. *Brain, 107, 1033-1072.*

Welch, L. (1989). The perception of moving plaids reveals two motion-processing stages. *Nature, 337, 734-736.*

Wilson, F.A.W., & Rolls E.T. (1993). The effects of stimulus novelty and familiarity on neuronal activity in the amygdala of monkeys performing recognition memory tasks. *Experimental Brain Research, 93, 367-382.*

Wong-Riley, M.T.T. (1976). Projections form the dorsal lateral geniculate nucleus to prestriate cortex in the squirrel monkey as demonstrated by retrograde transport of horseradish peroxidase. *Brain Research, 109, 595-600.*

Worgotter, F., & Holt, G. (1991). Spatiotemporal mechanisms in receptive fields of visual cortical simple cells: A model. *Journal of Neurophysiology, 65, 494-510.*

Worgotter, F., & Koch, C. (1991). A detailed model of the primary visual pathway in the cat: Comparison of afferent excitatory and intracortical inhibitory connection schemes for orientation selectivity. *Journal of Neuroscience, 11, 1959-1979.*

Yamaguchi, Y., & Shimizu, H. (1994). Pattern recognition with figure-ground seperation by generation of coherent oscillations. *Neural Networks, 7, 49-63.*

Yamane, S., Kaji, S., & Kawano, K. (1988). What facial features activate face neurons in the inferotemporal cortex of the monkey. *Experimental Brain Research, 73, 209-214.*

Yeterian E.H., & van Hosen G.W. (1978). Cortico-striate projections in the rhesus monkey: The organization of certain cortico-caudate connections. *Brain Research, 139, 43-63.*

Young, M.P. (1992). Objective analysis of the topological organization of the primate cortical visual system. *Nature, 358, 152-155.*

Young, M.P. (1993). Modules for pattern-recognition. *Current Biology, 3, 44-46.*

Young, M.P., & Yamane, S. (1992). Sparse population coding of faces in the inferotemporal cortex. *Science, 256, 1327-1331.*

Young, M.P., & Yamane, S. (1993). An analysis at the population level of the processing of faces in the inferotemporal cortex. In (T. Ono, L.R.Squire, R.E. Raichle, D.I. Perrett & M. Fukuda, Eds) *Brain Mechanisms of Perception and Memory: From Neuron to Behaviour*, pp 47-70. Oxford University Press, Oxford, England.

Young, M.P., Tanaka, K., & Yamane, S. (1992). On oscillating neuronal responses in the visual cortex of the monkey. *Journal of Neurophysiology, 67, 1464-1474.*

Yukie, M., & Iwai, E. (1981). Direct projection from the dorsal lateral geniculate nuclues to the prestriate cortex in macaque monkeys. *Journal of Comparative Neurology, 201, 81-97.*

Zeki, S., & Shipp, S. (1988). The functional logic of cortical connections. *Nature, 335, 311-317.*

Zeki, S., Watson, J.D.G., & Frackowiak, R.S.J. (1993). Going beyond the information given: The relation of illusory visual motion to brain activity. *Proceedings of the Royal Society of London (Series B), 252, 215-222.*

Zeki, S., Watson, J.D.G., Lueck, C.J., Friston, K.J., and Kennard, C.F. (1991). A direct demonstration of functional specialization in human visual cortex. *Journal of Neuroscience, 11, 641-649.*

Zeki, S.M. (1973). Colour coding in rhesus monkey prestriate cortex. *Brain Research, 53, 422-427.*

Zeki, S.M. (1974a) Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *Journal of Physiology (London), 236, 549-573.*

Zeki, S.M. (1974b). The mosaic organization of visual cortex in the monkey. In R. Bellairs and E.G. Gray (Eds), *Essays on the Nervous System: A Festschrift for Professor J.Z. Young* (pp 327-343). Oxford: Clarendon Press.

Zeki, S.M. (1977). Colour coding in the superior temporal sulcus of rhesus monkey visual cortex. *Journal of Physiology (London), 197, 195-223.*

Zeki, S.M. (1978). The third visual complex of rhesus monkey prestriate cortex. *Journal of Physiology (London), 277, 245-272.*

Zeki, S.M. (1978). Uniformity and diversity of structure and function in rhesus monkey prestriate visual cortex. *Journal of Physiology (London), 277, 273-290.*

Zeki, S.M. (1980). The representation of colours in the cerebral cortex. *Nature, 284, 412-418.*

Zeki, S.M. (1990). A theory of multistage integration in the visual cortex. In J.C. Eccles and O. Creutzfeldt (Eds) *The Principles of Design and Operation of the Brain, Experimental Brain Research Series* (Vol. 21, pp 963-971), Germany: Springer Verlag.

Zihl, J.D., Von Cramon, D., & Mai, N. (1983). Selective disturbance of vision after bilateral brain damage. *Brain, 106, 313-340.*